United States
Environmental Protection
Agency

Office of Research and
Development
Washington DC 20460

Office of Water
4305
Washington DC 20460

EPA/600/R-99/064
March 2000

EPA

Methods for Measuring the Toxicity and Bioaccumulation of Sediment-associated Contaminants with Freshwater Invertebrates

Second Edition

EPA 600/R-99/064
MARCH 2000

Methods for Measuring the Toxicity and Bioaccumulation of Sediment-associated Contaminants with Freshwater Invertebrates

Second Edition

Office of Research and Development
Mid-Continent Ecology Division
U.S. Environmental Protection Agency
Duluth, Minnesota 55804

Office of Science and Technology
Office of Water
U.S. Environmental Protection Agency
Washington, D.C. 20460

Disclaimer

This guidance is designed to describe procedures for testing freshwater organisms in the laboratory to evaluate the potential toxicity or bioaccumulation of chemicals in whole sediments. This guidance document has no immediate or direct regulatory consequence. It does not in itself establish or affect legal rights or obligations, or represent a determination of any party's liability. The USEPA may change this guidance in the future.

This guidance document has been reviewed in accordance with USEPA Policy and approved for publication. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

Foreword

Sediment contamination is a widespread environmental problem that can potentially pose a threat to a variety of aquatic ecosystems. Sediment functions as a reservoir for common chemicals such as pesticides, herbicides, polychlorinated biphenyls (PCBs), polycyclic aromatic hydrocarbons (PAHs), and metals such as lead, mercury, and arsenic. In-place contaminated sediment can result in depauperate benthic communities, while disposal of contaminated dredged material can potentially exert adverse effects on both pelagic and benthic systems. Historically, assessment of sediment quality has been limited to chemical characterizations. The United States Environmental Protection Agency (USEPA) is developing methodologies to calculate chemical-specific sediment quality guidelines (referred to as equilibrium partitioning sediment guidelines or ESGs) for use in the Agency's regulatory programs. However, quantifying contaminant concentrations alone cannot always provide enough information to adequately evaluate potential adverse effects that arise from interactions among chemicals, or that result from time-dependent availability of sediment-associated contaminants to aquatic organisms. Because relationships between bioavailability and concentrations of chemicals in sediment are not fully understood, determination of contaminated sediment effects on aquatic organisms may require the use of controlled toxicity and bioaccumulation tests.

As part of USEPA's Contaminated Sediment Management Strategy, Agency programs have agreed to use consistent methods to determine whether sediments have the potential to affect aquatic ecosystems. More than ten federal statutes provide authority to many USEPA program offices to address the problem of contaminated sediment. The sediment test methods in this manual will be used by USEPA to make decisions under a range of statutory authorities concerning such issues as: dredged material disposal, registration of pesticides, assessment of new and existing industrial chemicals, Superfund site assessment, and assessment and cleanup of hazardous waste treatment, storage, and disposal facilities. The use of uniform sediment testing procedures by USEPA programs is expected to increase data accuracy and precision, facilitate test replication, increase the comparative value of test results, and ultimately increase the efficiency of regulatory processes requiring sediment tests.

This second edition of the manual is a revision to USEPA (1994a; EPA 600/R-94/024). Primary revisions to the first edition of the manual include:

Section 14: This new section describes methods for evaluating sublethal effects of sediment-associated contaminants with the amphipod *Hyalella azteca*. See also associated revisions to Sections 1.3, 2, 4.3, 7.1.3, and 10.3. Section 11 also outlines methods for measuring growth and survival as primary endpoints in 10-d tests with *Hyalella azteca.*

Section 15: This new section describes methods for evaluating sublethal effects of sediment-associated contaminants with the midge *Chironomus tentans*. See also associated revisions to Sections 1.3, 2, 4.3, 7.1.3, 10.4, and Appendix C.

Section 2.1.2.1.1: Additional detail has been included on test acceptability (i.e., control vs. reference sediment).

Foreword (continued)

Section 6.2.2: The range of acceptable light intensity for culture and testing has been revised from 500 lux to 1000 lux to 100 to 1000 lux.

Sections 7.2, 8.2, 8.3.2, 8.4.4.7: Additional detail has been added to sections on formulated sediments, sediment storage, sediment spiking, and interstitial water sampling.

Sections 9.14, 10.3, and 17.4: The requirement to conduct monthly reference-toxicity tests has been modified to recommend the conduct of reference-toxicity tests periodically to assess the sensitivity of the test organisms.

Sections 9.14.2 and 17.4.3: These revised sections now state that before conducting tests with contaminated sediment, it is strongly recommended that the laboratory conduct the tests with control sediment(s). Results of these preliminary studies should be used to determine if use of the control sediment and other test conditions (i.e., water quality) result in acceptable performance in the tests as outlined in Tables 11.3, 12.3, 13.4, 14.3, and 15.3.

Section 10.3.2: Diatoms are no longer used to culture *Hyalella azteca* following procedures of USEPA (1993).

Section 11: *In Section11.2.2 (and associated sections and tables):* The recommended feeding level of 1.5 mL of YCT/day/beaker in the 10-d *Hyalella azteca* sediment toxicity test in the first edition of the manual has been revised to 1.0 ml of YCT/day/beaker. This change was made to make the 10-d test described in Section 11 consistent with the feeding level recommended in the 42-d test with *Hyalella azteca* described in Section 14. *In Section 11.3:* Additional guidance has been included in the revised manual regarding acclimation of test organisms to temperature (see also Section 12.3, 13.3, 14.3, and 15.3). *In Section 11.3.6.1.1:* Acceptable concentrations of dissolved oxygen in overlying water are now expressed in mg/L rather than in a percentage of saturation. See also Sections 10, 12, 13, 14, and 15.

Sections 12.3.8 and 15.3.8: The recommendation is now made to measure ash-free dry weight of *Chironomus tentans* instead of dry weight. See also Sections 13.3.8 for *Lumbriculus variegatus* and 14.3.7 for *Hyalella azteca*.

Section 13.3.7: This section outlines additional guidance on depuration of *Lumbriculus variegatus* in bioaccumulation testing.

Section 17.6: This revised section now includes summaries of the results of round-robin tests using the methods for long-term toxicity tests outlined in Sections 14 and 15.

Appendix A in the first edition of the manual (USEPA, 1994) was not included in this edition (summary of a workshop designed to develop consensus for the 10-d toxicity test and bioaccumulation methods). This information has been cited by reference in this current edition of the manual.

For additional guidance on the technical considerations in the manual, please contact Teresa Norberg-King, USEPA, Duluth, MN (218/529-5163, fax -5003, email norberg-king.teresa@epa.gov) or Chris Ingersoll, USGS, Columbia, MO (573/876-1819, fax -1896, email chris_ingersoll@usgs.gov).

Abstract

Procedures are described for testing freshwater organisms in the laboratory to evaluate the potential toxicity or bioaccumulation of chemicals in whole sediments. Sediments may be collected from the field or spiked with compounds in the laboratory. Toxicity methods are outlined for two organisms, the amphipod *Hyalella azteca* and the midge *Chironomus tentans.* Toxicity tests with amphipods or midges are conducted for 10 d in 300-mL chambers containing 100 mL of sediment and 175 mL of overlying water. Overlying water is renewed daily and test organisms are fed during the toxicity tests. The endpoints in the 10-d toxicity test with *H. azteca* and *C. tentans* are survival and growth. Procedures are primarily described for testing freshwater sediments; however, estuarine sediments (up to 15‰ salinity) can also be tested in 10-d sediment toxicity tests with *H. azteca.* Guidance is also provided for conducting long-term sediment toxicity tests with *H. azteca* and *C. tentans.* The long-term sediment exposures with *H. azteca* are started with 7- to 8-d-old amphipods. On Day 28 of the sediment exposure, amphipods are isolated from the sediment and placed in water-only chambers where reproduction is measured on Day 35 and 42. Endpoints measured in the amphipod test include survival (Day 28, 35, and 42), growth (on Day 28 and 42), and reproduction (number of young/female produced from Day 28 to 42). The long-term sediment exposures with *C. tentans* start with newly hatched larvae (<24-h old) and continue through emergence, reproduction, and hatching of the F_1 generation (about 60-d sediment exposures). Survival and growth are determined at 20 d. Starting on Day 23 to the end of the test, emergence and reproduction of *C. tentans* are monitored daily. The number of eggs/female is determined for each egg mass, which is incubated for 6 d to determine hatching success. The procedures described in Sections 14 and 15 include measurement of a variety of lethal and sublethal endpoints with *Hyalella azteca* and *Chironomus tentans*; minor modifications of the basic methods can be used in cases where only a subset of these endpoints is of interest. Guidance for conducting 28-d bioaccumulation tests with the oligochaete *Lumbriculus variegatus* is also provided in the manual. Overlying water is renewed daily and test organisms are not fed during bioaccumulation tests. Methods are also described for determining bioaccumulation kinetics of different classes of compounds during 28-d exposures with *L. variegatus.*

Contents

Foreword iii
Abstract v
Acknowledgments xvii

1 Introduction 1
1.1 Significance of Use 1
1.2 Program Applicability 2
1.3 Scope and Application 4
1.4 Performance-based Criteria 10

2 Summary of Method 11
2.1 Method Description and Experimental Design 11
2.2 Types of Tests 13
2.3 Test Endpoints 13

3 Definitions 14
3.1 Terms 14

4 Interferences 16
4.1 General Introduction 16
4.2 Noncontaminant Factors 17
4.3 Changes in Bioavailability 18
4.4 Presence of Indigenous Organisms 18

5 Health, Safety, and Waste Management 19
5.1 General Precautions 19
5.2 Safety Equipment 19
5.3 General Laboratory and Field Operations 19
5.4 Disease Prevention 20
5.5 Safety Manuals 20
5.6 Pollution Prevention, Waste Management, and Sample Disposal 20

6 Facilities, Equipment, and Supplies 21
6.1 General 21
6.2 Facilities 21
6.3 Equipment and Supplies 21

Contents (continued)

7 Water, Formulated Sediment, Reagents, and Standards 24
7.1 Water 24
7.2 Formulated Sediment 25
7.3 Reagents 28
7.4 Standards 28

8 Sample Collection, Storage, Manipulation, and Characterization 29
8.1 Collection 29
8.2 Storage 29
8.3 Manipulation 30
8.4 Characterization 31

9 Quality Assurance and Quality Control 33
9.1 Introduction 33
9.2 Performance-based Criteria 33
9.3 Facilities, Equipment, and Test Chambers 33
9.4 Test Organisms 34
9.5 Water 34
9.6 Sample Collection and Storage 34
9.7 Test Conditions 34
9.8 Quality of Test Organisms 34
9.9 Quality of Food 34
9.10 Test Acceptability 34
9.11 Analytical Methods 34
9.12 Calibration and Standardization 34
9.13 Replication and Test Sensitivity 35
9.14 Demonstrating Acceptable Performance 35
9.15 Documenting Ongoing Laboratory Performance 35
9.16 Reference Toxicants 35
9.17 Record Keeping 36

10 Collecting, Culturing, and Maintaining Test Organisms 38
10.1 Life Histories 38
10.2 General Culturing Procedures 40
10.3 Culturing Procedures for *Hyalella azteca* 41
10.4 Culturing Procedures for *Chironomus tentans* 42
10.5 Culturing Procedures for *Lumbriculus variegatus* 46

Contents (continued)

11 Test Method 100.1: *Hyalella azteca* 10-d Survival and Growth Test for Sediments 47

11.1 Introduction 47

11.2 Recommended Test Method for Conducting a 10-d Sediment Toxicity Test with *Hyalella azteca* 47

11.3 General Procedures 47

11.4 Interpretation of Results 52

12 Test Method 100.2: *Chironomus tentans* 10-d Survival and Growth Test for Sediments 55

12.1 Introduction 55

12.2 Recommended Test Method for Conducting a 10-d Sediment Toxicity Test with *Chironomus tentans* 55

12.3 General Procedures 55

12.4 Interpretation of Results 60

13 Test Method 100.3: *Lumbriculus variegatus* Bioaccumulation Test for Sediments 63

13.1 Introduction 63

13.2 Procedure for Conducting Sediment Bioaccumulation Tests with *Lumbriculus variegatus* 63

13.3 General Procedures 64

13.4 Interpretation of Results 71

14 Test Method 100.4: *Hyalella azteca* 42-d Test for Measuring the Effects of Sediment-associated Contaminants on Survival, Growth, and Reproduction 72

14.1 Introduction 72

14.2 Procedure for Conducting a *Hyalella azteca* 42-d Test for Measuring the Effects of Sediment-associated Contaminants on Survival, Growth, and Reproduction 74

14.3 General Procedures 75

14.4 Interpretation of Results 78

15 Test Method 100.5: Life-cycle Test for Measuring the Effects of Sediment-associated Contaminants on *Chironomus tentans* 84

15.1 Introduction 84

15.2 Procedure for Conducting a Life-cycle Test for Measuring the Effects of Sediment-associated Contaminants on *Chironomus tentans* 84

15.3 General Procedures 87

15.4 Interpretation of Results 92

16 Data Recording, Data Analysis and Calculations, and Reporting 97

16.1 Data Recording 97

16.2 Data Analysis 97

16.3 Data Interpretation 113

16.4 Reporting 114

Contents (continued)

17 Precision and Accuracy 115
17.1 Determining Precision and Accuracy 115
17.2 Accuracy 115
17.3 Replication and Test Sensitivity 116
17.4 Demonstrating Acceptable Laboratory Performance 116
17.5 Precision of Sediment Toxicity Test Methods: Evaluation of 10-d Sediment Tests and Reference-toxicity Tests 117
17.6 Precision of Sediment Toxicity Test Methods: Evaluation of Long-term Sediment Tests 127
18 References 141

Appendices

A. Exposure Systems 157
B. Food Preparation 169
B.1 Yeast, Cerophyl®, and Trout Chow (YCT) for Feeding the Cultures and *Hyalella azteca* 169
B.2 Algal Food 170
B.3 Tetrafin® Food (or Other Fish Flake Food) for Culturing and Testing *Chironomus tentans* 172
C. Supplies and Equipment for Conducting the *Chironomus tentans* Long-term Sediment Toxicity Test 173
C.1 General 173
C.2 Emergence Traps 173
C.3 Reproduction/Oviposit Chambers 173
C.4 Adult Collector Dish 173
C.5 Aspirator 173
C.6 Auxiliary Male Holding Dish 173
C.7 Egg Hatching Chamber 173
C.8 Supplies and Sources 174
C.9 Construction of an Adult Midge Emergence Trap for Use in a "Zumwalt" Exposure System in Life-cycle Sediment Tests 175
D. Sample Data Sheets 177

Figures

Figure 10.1 Mean length (+/- 2SD) and relative age of *Hyalella azteca* collected by sieving in comparison with length of known-age organisms. 43

Figure 10.2 *Chironomus tentans* larvae 43

Figure 10.3 Aspirator chamber (A) and reproduction and oviposit chamber (B) for adult midges. 45

Figure 11.1 *Hyalella azteca*. 51

Figure 11.2 Lifestage sensitivity of *Hyalella azteca* in 96-h water-only exposures. 53

Figure 11.3 Average recovery of different age *Hyalella azteca* from sediment by 7 individuals. 54

Figure 12.1 Lifestage sensitivity of chironomids. 61

Figure 13.1 Predicted depuration of nonionic organic chemicals from tissue of *Lumbriculus variegatus* as a function of K_{OW} and duration of depuration, assuming no contribution of sediment in the gut. 69

Figure 14.1 Relationships between *Hyalella azteca* length and reproduction 80

Figure 14.2 Relationships between *Hyalella azteca* dry weight and reproduction. 81

Figure 14.3 Relationship between *Hyalella azteca* length and dry weight. 82

Figure 15.1 Relationship between weight and emergence of *Chironomus tentans* 94

Figure 15.2 Relationship between weight and reproduction of *Chironomus tentans* 94

Figure 15.3 Relationship between ash-free dry weight (AFDW) and length of *Chironomus tentans* 95

Figure 15.4 Relationship between ash-free dry weight (AFDW) and intrinsic rate of natural increase of *Chironomus tentans* 95

Figure 16.1 Treatment response for a Type I and Type II error. 100

Figure 16.2 Power of the test vs. percent reduction in treatment response relative to the control mean at various CVs (8 replicates, alpha = 0.05 [one-tailed]) 101

Figure 16.3 Power of the test vs. percent reduction in treatment response relative to the control mean at various CVs (5 replicates, alpha = 0.05 [one-tailed]). 101

Figure 16.4 Power of the test vs. percent reduction in treatment response relative to the control mean at various CVs (8 replicates, alpha = 0.10 [one-tailed]). 102

Figure 16.5 Effect of CV and number of replicates on the power to detect a 20% decrease in treatment response relative to the control mean (alpha = 0.05 [one-tailed]). 102

Figure 16.6 Effect of alpha and beta on the number of replicates at various CVs (assuming combined alpha + beta = 0.25) 103

Figure 16.7 Decision tree for analysis of survival, growth, and reproduction data subjected to hypothesis testing 104

Figure 16.8 Decision tree for analysis of point estimate data 108

Figures (continued)

Figure 17.1 Control (cusum) charts: (A) hypothesis testing results; and (B) point estimates (LC, EC, or IC) 116

Figure A.1 Portable table top STIR system described in Benoit et al. (1993). 158

Figure A.2 Portable table top STIR system with several additional options as described in Benoit et al. (1993). 159

Figure A.3 Tanks for the STIR system in Benoit et al. (1993) 160

Figure A.4 Water splitting chamber described in Zumwalt et al. (1994) 165

Figure A.5 Diagram of in-line flow splitter used to deliver overlying water in the sediment exposures of *Lumbriculus variegatus* (Brunson et al., 1998) 167

Figure C.1 Emergence trap used in the life-cycle *Chironomus tentans* sediment test 174

Figure C.2 The reproduction/oviposit chamber with the double stack support stand 174

Figure C.3 Adult collection/transfer equipment 175

Figure C.4 Emergence traps that can be used with the Zumwalt water-delivery system described in Section A.4 176

Figure D.1 Data sheet for the evaluation of a *Chironomus tentans* culture 178

Figure D.2 QA/QC data sheet for *Chironomus tentans* culture 179

Figure D.3 QA/QC data sheets for *Chironomus tentans* culture 180

Figure D.4 Data sheet for performing reference-toxicity tests 181

Figure D.5 Data sheet for temperature and overlying water chemistry measurements 182

Figure D.6 Data sheet for daily checklist for sediment tests 183

Figure D.7 Data sheet for water quality parameters 184

Figure D.8 Chemistry data sheet 185

Figure D.9 Daily comment data sheet 186

Figure D.10 Weight data sheet 187

Figure D.11 Data sheets for *Chironomus tentans* tests 188

Figure D.12 Instructions for terminating a *Chironomus tentans* test 189

Figure D.13 Data sheet for the *Chironomus tentans* life-cycle test 190

Figure D.14 Example entries for a *Chironomus tentans* life-cycle test data sheet 191

Figure D.15 Instructions for completing the *Chironomus tentans* life-cycle test data sheet 192

Tables

Table 1.1 Sediment Quality Assessment Procedures 3

Table 1.2 Statutory Needs for Sediment Quality Assessment 4

Table 1.3 Rating of Selection Criteria for Freshwater Sediment Toxicity Testing Organisms . 7

Table 1.4 Water-only, 10-d LC50 (µg/L) Values for *Hyalella azteca, Chironomus tentans,* and *Lumbriculus variegatus* 7

Table 4.1 Advantages and Disadvantages for Use of Sediment Tests 16

Table 6.1 Equipment and Supplies for Culturing and Testing Specific Test Organisms 23

Table 7.1 Characteristics of Three Sources of Clays and Silts Used in Formulated Sediments 26

Table 7.2 Carbon, Nitrogen, Phosphorus Levels for Various Sources of Organic Carbon.... 26

Table 7.3 Sources of Components Used in Formulated Sediments 27

Table 9.1 Recommended Test Conditions for Conducting Reference-toxicity Tests with One Organism/Chamber 36

Table 9.2 Recommended Test Conditions for Conducting Reference-toxicity Tests with More Than One Organism/Chamber 37

Table 10.1 Sources of Starter Cultures of Test Organisms 40

Table 10.2 *Chironomus tentans* Instar and Head Capsule Widths 43

Table 11.1 Test Conditions for Conducting a 10-d Sediment Toxicity Test with *Hyalella azteca* 48

Table 11.2 General Activity Schedule for Conducting a 10-d Sediment Toxicity Test with *Hyalella azteca* 48

Table 11.3 Test Acceptability Requirements for a 10-d Sediment Toxicity Test with *Hyalella azteca* 49

Table 12.1 Recommended Test Conditions for Conducting a 10-d Sediment Toxicity Test with *Chironomus tentans* 56

Table 12.2 General Activity Schedule for Conducting a 10-d Sediment Toxicity Test with *Chironomus tentans* 57

Table 12.3 Test Acceptability Requirements for a 10-d Sediment Toxicity Test with *Chironomus tentans* 58

Table 13.1 Recommended Test Conditions for Conducting a 28-d Sediment Bioaccumulation Test with *Lumbriculus variegatus* 64

Table 13.2 Recommended Test Conditions for Conducting a Preliminary 4-d Sediment Toxicity Screening Test with *Lumbriculus variegatus* 65

Tables (continued)

Table 13.3 General Activity Schedule for Conducting a 28-d Sediment Bioaccumulation Test with *Lumbriculus variegatus* 66

Table 13.4 Test Acceptability Requirements for a 28-d Sediment Bioaccumulation Test with *Lumbriculus variegatus* 67

Table 13.5 Grams of *Lumbriculus variegatus* Tissue (Wet Weight) Required for Various Analytes at Selected Lower Limits of Detection 70

Table 13.6 Detection Limits (ng) of Individual PAHs by HPLC-FD 70

Table 14.1 Test Conditions for Conducting a 42-d Sediment Toxicity Test with *Hyalella azteca* 73

Table 14.2 General Activity Schedule for Conducting a 42-d Sediment Toxicity Test with *Hyalella azteca* 74

Table 14.3 Test Acceptability Requirements for a 42-d Sediment Toxicity Test with *Hyalella azteca* 75

Table 14.4 Percentage of Paired Tests or Paired Endpoints Identifying Samples as Toxic in *Hyalella azteca* 14-d or 28-d Tests 83

Table 15.1 Test Conditions for Conducting a Long-term Sediment Toxicity Test with *Chironomus tentans* 85

Table 15.2 General Activity Schedule for Conducting a Long-term Sediment Toxicity Test with *Chironomus tentans* 86

Table 15.3 Test Acceptability Requirements for a Long-term Sediment Toxicity Test with *Chironomus tentans* 87

Table 15.4 Endpoints for a Long-term Sediment Toxicity Test with *Chironomus tentans* 88

Table 16.1 Suggested α Levels to Use for Tests of Assumptions 105

Table 16.2 Estimated Time to Obtain 95 Percent of Steady-state Tissue Residue 112

Table 17.1 Intralaboratory Precision for Survival of *Hyalella azteca* and *Chironomus tentans* in 10-d Whole-sediment Toxicity Tests, June 1993 117

Table 17.2 Participants in 1993 Round-robin Studies 118

Table 17.3 Interlaboratory Precision for *Hyalella azteca* 96-h LC50s from Water-only Static Acute Toxicity Tests Using a Reference Toxicant (KCl) (October 1992) .. 118

Table 17.4 Interlaboratory Precision for Survival of *Hyalella azteca* in 10-d Whole-sediment Toxicity Tests Using Four Sediments (March 1993) 119

Table 17.5 Interlaboratory Precision for *Chironomus tentans* 96-h LC50s from Water-only Static Acute Toxicity Tests Using a Reference Toxicant (KCl) (December 1992) 120

Table 17.6 Interlaboratory Precision for *Chironomus tentans* 96-h LC50s from Water-only Static Acute Toxicity Tests Using a Reference Toxicant (KCl) (May 1993)) 120

Table 17.7 Interlaboratory Precision for Survival of *Chironomus tentans* in 10-d Whole-sediment Toxicity Tests Using Three Sediments (May 1993) 121

Tables (continued)

Table 17.8 Interlaboratory Precision for Growth of *Chironomus tentans* in 10-d Whole-sediment Toxicity Tests Using Three Sediments (May 1993) 121

Table 17.9 Interlaboratory Precision for Survival (%) of *Hyalella azteca* in 10-d Whole-sediment Toxicity Tests (1996/1997) 123

Table 17.10 Interlaboratory Precision for Survival (%) of *Chironomus tentans* in 10-d Whole-sediment Toxicity Tests (1996/1997) 124

Table 17.11 Interlaboratory Precision for Growth (mg/Individual dry weight) of *Chironomus tentans* in 10-d Whole-sediment Toxicity Tests (1996/1997) 125

Table 17.12 Interlaboratory Precision for Growth (mg/Individual as ash-free dry weight) of *Chironomus tentans* in 10-d Whole-sediment Toxicity Tests (1996/1997) 126

Table 17.13 Physical Characteristics of the Sediments Used in the Preliminary and Definitive Round-robin Evaluations of Long-term Methods for Sediment Toxicity Testing (Section 17.6) 128

Table 17.14 Percentage of Laboratories Meeting Performance Levels for the Following Endpoints in the WB Control Sediment Evaluated in the Long-term Round-robin Tests 128

Table 17.15 Interlaboratory Comparison of Day 28 Percent Survival (Mean ±SD) of *H. azteca* in a Long-term Sediment Exposure Using Five Sediments 129

Table 17.16 Interlaboratory Comparison of Day 35 Percent Survival (Mean ±SD) of *H. azteca* in a Long-term Sediment Exposure Using Five Sediments 130

Table 17.17 Interlaboratory Comparison of Day 42 Percent Survival (Mean ±SD) of *H. azteca* in a Long-term Sediment Exposure Using Five Sediments 131

Table 17.18 Interlaboratory Comparison of Day 28 Length (Mean mm/Individual ±SD) of *H. azteca* in a Long-term Sediment Exposure Using Five Sediments 132

Table 17.19 Interlaboratory Comparison of Day 28 Dry Weight (Mean mg/Individual ±SD) of *H. azteca* in a Long-term Sediment Exposure Using Five Sediments 133

Table 17.20 Interlaboratory Comparison of Reproduction (Mean Number of Young/Female ±SD) of *H. azteca* in a Long-term Sediment Exposure Using Five Sediments ... 134

Table 17.21 Interlaboratory Comparison of Day 20 Percent Survival (Mean ±SD) of *C. tentans* in a Long-term Sediment Exposure Using Five Sediments 135

Table 17.22 Interlaboratory Comparison of Dry Weight (Mean mg/Individual ±SD) of *C. tentans* in a Long-term Sediment Exposure Using Five Sediments 136

Table 17.23 Interlaboratory Comparison of Ash-free Dry Weight (Mean mg/Individual ±SD) of *C. tentans* in a Long-term Sediment Exposure Using Five Sediments 137

Tables (continued)

Table 17.24 Interlaboratory Comparison of Percent Emergence (Mean ±SD) of *C. tentans* in a Long-term Sediment Exposure Using Five Sediments 138

Table 17.25 Interlaboratory Comparison of the Number of Eggs/Female (Mean ±SD) in a Long-term Sediment Exposure Using Five Sediments 139

Table 17.26 Interlaboratory Comparison of Percent Hatch (Mean ±SD) of *C. tentans* in a Long-term Sediment Exposure Using Five Sediments 140

Table A.1 Sediment Copper Concentrations and Organism Survival and Growth at the End of a 10-d Test 163

Table A.2 Sediment Dieldrin Concentrations and Organism Survival and Growth at the End of a 10-d Test 163

Table A.3 Materials Needed for Constructing a Zumwalt et al. (1994) Delivery System 166

Table B.1 Nutrient Stock Solutions for Maintaining Algal Stock Cultures 170

Table B.2 Final Concentration of Macronutrients and Micronutrients in the Algal Culture Medium 170

Acknowledgments

This document is a general purpose testing manual for freshwater sediments. This manual is a revision to a previously published edition of this manual (USEPA, 1994a). The approaches described in this manual were developed from ASTM (1999a), ASTM (1999b), ASTM (1999c), ASTM (1999d), Ankley et al. (1993), Phipps et al. (1993), USEPA (1994b), USEPA (1994c), Ingersoll et al. (1995), Ingersoll et al. (1998), Sibley et al. (1996), Sibley et al. (1997a), Sibley et al. (1997b), and Benoit et al. (1997).

This second edition of the manual reflects the consensus of the Freshwater Sediment Toxicity Assessment Committee and the U.S. Environmental Protection Agency (USEPA) Program Offices. Members of the Freshwater Sediment Toxicity Assessment Committee for the second edition of this manual included G.A. Burton, Wright State University, Dayton, OH; T.D. Dawson, Integrated Laboratory Systems (ILS), Duluth, MN; F.J. Dwyer, U.S. Fish & Wildlife Service (USFWS), Columbia, MO; C.G. Ingersoll, U.S. Geological Survey (USGS), Columbia, MO; D.S. Ireland, USEPA, Washington, D.C.; N.E. Kemble, USGS, Columbia, MO; D.R. Mount, USEPA, Duluth, MN; T.J. Norberg-King, USEPA, Duluth, MN; P.K. Sibley, University of Guelph, Guelph, Ontario, and Leanne Stahl, USEPA, Washington, D.C.

The principal authors of the first edition of this manual (USEPA, 1994a) included C.G. Ingersoll, G.T. Ankley, G.A. Burton, F.J. Dwyer, R.A. Hoke, T.J. Norberg-King, and P.V. Winger. Principal authors to the second edition of the manual included C.G. Ingersoll, G.A. Burton, T.D. Dawson, F.J. Dwyer, D.S. Ireland, N.E. Kemble, D.R. Mount, T.J. Norberg-King, P.K. Sibley, and L. Stahl.

Contributors to specific sections of the manual are:

1. Sections 1-9; General Guidelines

   G.T. Ankley, USEPA, Duluth, MN
   G.A. Burton, Wright State University, Dayton, OH
   T.D. Dawson, ILS, Duluth, MN
   F.J. Dwyer, USGS, Columbia, MO
   R.A. Hoke, SAIC, Hackensack, NJ
   C.G. Ingersoll, USGS, Columbia, MO
   D.S. Ireland, USEPA, Washington, DC
   N.E. Kemble, USGS, Columbia, MO
   D.R. Mount, USEPA, Duluth, MN
   T.J. Norberg-King, USEPA, Duluth, MN
   C.E. Schlekat, SAIC, Narragansett, RI
   K.J. Scott, SAIC, Narragansett, RI
   L. Stahl, USEPA, Washington, DC

2. Sections 10-15; Culture and Test Methods

   G.T. Ankley, USEPA, Duluth, MN
   D.A. Benoit, USEPA, Duluth, MN
   T.D. Dawson, ILS, Duluth, MN
   E.L. Brunson, USGS, Columbia, MO

Acknowledgements (continued)

F.J. Dwyer, USGS, Columbia, MO
I.E. Greer, USGS, Columbia, MO
R.A. Hoke, SAIC, Hackensack, NJ
C.G. Ingersoll, USGS, Columbia, MO
N.E. Kemble, USGS, Columbia, MO
P.F. Landrum, NOAA, Ann Arbor, MI
H. Lee, USEPA, Newport, OR
D.R. Mount, USEPA, Duluth, MN
T.J. Norberg-King, USEPA, Duluth, MN
P.K. Sibley, University of Guelph, Guelph, Ontario
P.V. Winger, USGS, Athens, GA

3. Section 16; Statistical Analysis

J. Heltshe, SAIC, Narragansett, RI
R.A. Hoke, SAIC, Hackensack, NJ
H. Lee, USEPA, Newport, OR
T.J. Norberg-King, USEPA, Duluth, MN
C. Schlekat, SAIC, Narragansett, RI

4. Section 17; Precision and Accuracy

G.T. Ankley, USEPA, Duluth, MN
G.A. Burton, Wright State University, Dayton, OH
M.S. Greenburg, Wright State University, Dayton, OH
C.G. Ingersoll, USGS, Columbia, MO
N.E. Kemble, USGS, Columbia, MO
T.J. Norberg-King, USEPA, Duluth, MN
C.D. Rowland, Wright State University, Dayton, OH
P.K. Sibley, University of Guelph, Guelph, Ontario

Review comments from the following individuals are gratefully acknowledged on the first edition of this manual (USEPA, 1994a): T. Armitage, USEPA, OST, Washington, D.C.; J. Arthur, R. Spehar, and C. Stephan, USEPA, ORD, Duluth, MN; T. Bailey, USEPA, OPP, Washington, D.C.; C. Philbrick Barr and P. Nolan, USEPA, Region 1, Lexington, MA; S. Collyard, T. Dawson, J. Jenson, J. Juenemann, and J. Thompson, ILS, Duluth, MN; P. Crocker and S. McKinney, USEPA, Region 6, Dallas, TX; S. Ferraro and R. Swartz, ORD, Newport, OR; L. Gast, TAI, Newtown, OH; G. Hanson, USEPA, OSW, Washington, D.C.; D. Klemm, EMSL, Newtown, OH; D. Reed, USEPA, OWM, Washington, D.C.; C. Scheklat and J. Scott, SAIC, Narragansett, RI; F. Schmidt, USEPA, OWOW, Washington, D.C.; J. Smrchek, USEPA, OPPT, Washington, D.C.

Review comments for the following individuals are gratefully acknowledged on the second edition of the manual: P. Crocker, USEPA, Region 6, Dallas, TX; P. De Lisle, Coastal Bioanalysts, Gloucester Point, VA; R. Haley, NCASI, Anacortes, WA; R. Hoke, DuPont, Newark, DE; S. Kroner, USEPA, OSW, Washington, D.C.; P. Landrum, NOAA, Ann Arbor, MI; J. Lazorchak, USEPA, ORD, Cincinnati, OH; S. Lin, USEPA, OWOW, Washington, D.C.; A. Samel, DuPont, Newark, DE; J. Smrchek, USEPA, OPPT, Washington, D.C.; M. Thompson, USEPA, OST, Washington, D.C.; P. Winger, USGS, Athens, GA.

Acknowledgements (continued)

Participation by the following laboratories in the round-robin testing is greatly appreciated for the first edition of the manual (USEPA, 1994a): ABC Laboratories, Columbia, MO; Environment Canada, Burlington, Ontario; EVS Consultants, Vancouver, BC; Michigan State University, East Lansing, MI; National Fisheries Contaminant Research Center, Athens, GA; Midwest Science Center, Columbia, MO; Center University of Mississippi, University, MS; University of Wisconsin-Superior, Superior, WS; USEPA, Cincinnati, OH; USEPA, Duluth, MN; Washington Department of Ecology, Manchester, WA; Wright State University, Dayton, OH. Culturing support was supplied for USEPA Duluth by S. Collyard, J. Juenemann, J. Jenson, and J. Denny.

Participation by the following laboratories in the round-robin testing is greatly appreciated for this second edition of the manual: Aquatech Biological Sciences, South Burlington, VT; AScI Environment, Duluth, MN; Arkansas State University, State University, AR; Bayer, Stillwell, KS; Beak, Brampton, Ontario; Carolina Ecotox, Durham, NC; Columbia Environmental Research Center, Columbia, MO; Dell Engineering, Holland, MI; EA Engineering, Sparks, MD; Great Lakes Research Center, Traverse City, MI; Michigan State University, East Lansing, MI; NCASI, Anacortes, WA; Patuxent Wildlife Research Center, Athens, GA; SAIC, Naragansett, RI; Springborn Laboratories, Wareham, MA; University of Mississippi, Oxford, MS; Wright State University, Dayton, OH; USEPA, Duluth, MN; USEPA-Region 1, Lexington, MA; Zeneca, Bracknell, Berks, United Kingdom.

USEPA's Office of Science and Technology provided support for the development of this manual.

Section 1
Introduction

1.1 Significance of Use

1.1.1 Sediment provides habitat for many aquatic organisms and is a major repository for many of the more persistent chemicals that are introduced into surface waters. In the aquatic environment, most anthropogenic chemicals and waste materials including toxic organic and inorganic chemicals eventually accumulate in sediment. Mounting evidence exists of environmental degradation in areas where USEPA Water Quality Criteria (WQC; Stephan et al., 1985) are not exceeded, yet organisms in or near sediments are adversely affected (Chapman, 1989). The WQC were developed to protect organisms in the water column and were not intended to protect organisms in sediment. Concentrations of chemicals in sediment may be several orders of magnitude higher than in the overlying water; however, bulk sediment concentrations have not been strongly correlated to bioavailability (Burton, 1991). Partitioning or sorption of a compound between water and sediment may depend on many factors, including aqueous solubility, pH, redox, affinity for sediment organic carbon and dissolved organic carbon, grain size of the sediment, sediment mineral constituents (oxides of iron, manganese, and aluminum), and the quantity of acid volatile sulfides in sediment (Di Toro et al., 1990, 1991). Although certain chemicals are highly sorbed to sediment, these compounds may still be available to the biota. Contaminated sediments may be directly toxic to aquatic life or can be a source of contaminants for bioaccumulation in the food chain.

1.1.2 Assessments of sediment quality have commonly included sediment chemical analyses and surveys of benthic community structure. Determination of sediment chemical concentrations on a dry weight basis alone offers little insight into predicting adverse biological effects because bioavailability may be limited by the intricate partitioning factors mentioned above. Likewise, benthic community surveys may be inadequate because they sometimes fail to discriminate between effects of contaminants and those that result from unrelated non-contaminant factors, including water-quality fluctuations, physical parameters, and biotic interactions. To obtain a direct measure of sediment toxicity or bioaccumulation, laboratory tests have been developed in which surrogate organisms are exposed to sediments under controlled conditions. Sediment toxicity tests have evolved into effective tools that provide direct, quantifiable evidence of biological consequences of sediment contamination that can only be inferred from chemical or benthic community analyses. To evaluate sediment quality nationwide, USEPA developed the National Sediment Inventory (NSI), which is a compilation of existing sediment quality data and protocols used to evaluate the data. The NSI was used to produce the first biennial report to Congress on sediment quality in the United States as required under the Water Resources Development Act of 1992 (USEPA, 1997a; 1997b; 1997c). USEPA's evaluation of the data shows that sediment contamination exists in every region and state of the country and various waters throughout the United States contain sediment that is sufficiently contaminated with toxic pollutants to pose potential risks to fish and to humans and wildlife who eat fish. The use of consistent sediment testing methods described in this manual will provide high quality data needed for the NSI, future reports to Congress, and regulatory programs to prevent, remediate, and manage contaminated sediments (USEPA, 1998).

1.1.3 The objective of a sediment test is to determine whether chemicals in sediment are harmful to or are bioaccumulated by benthic organisms. The tests can be used to measure interactive toxic effects of complex chemical mixtures in sediment. Furthermore, knowledge of specific pathways of interactions among sediments and test organisms is not necessary to conduct the tests (Kemp and Swartz, 1988). Sediment tests can be used to (1) determine the relationship between toxic effects and bioavailability; (2) investigate interactions among chemicals; (3) compare the sensitivities of different organisms; (4) determine spatial and temporal distribution of contamination; (5) evaluate dredged material; (6) measure toxicity as part of product licensing or safety testing or chemical approval; (7) rank areas for cleanup, and (8) set cleanup goals and estimate the effectiveness of remediation or management practices.

1.1.4 A variety of standard methods have been developed for assessing the toxicity of contaminants associated with sediments using amphipods, midges, polychaetes, oligochaetes, mayflies, or cladocerans (i.e., ASTM,1999a; ASTM,1999b; ASTM, 1999c; ASTM, 1999d; USEPA, 1994a; USEPA, 1994b; Environment Canada, 1997a; Environment Canada, 1997b). Several endpoints are suggested in these methods to measure effects of contaminants in sediment including survival, growth, behavior, or reproduction; however, survival of test organisms in

10-d exposures is the endpoint most commonly reported. These short-term exposures which only measure effects on survival can be used to identify high levels of contamination, but may not be able to identify moderately contaminated sediments (Sibley et al., 1996; Sibley et al., 1997a; Sibley et al., 1998; Benoit et al., 1997; Ingersoll et al., 1998). Sublethal endpoints in sediment tests may also prove to be better estimates of responses of benthic communities to contaminants in the field (Kemble et al., 1994) The first edition of this manual (USEPA, 1994a) described 10-d toxicity tests with the amphipod *Hyalella azteca* and midge *Chironomus tentans* (Section 11, 12). This second edition of the manual now outlines approaches for evaluating sublethal endpoints in longer-term sediment exposures with these two species (Section 14, 15). Guidance is also presented in Section 13 regarding sediment bioaccumulation testing with the oligochaete *Lumbriculus variegatus*.

1.1.5 Results of toxicity tests on sediments spiked at different concentrations of chemicals can be used to establish cause and effect relationships between chemicals and biological responses. Results of toxicity tests with test materials spiked into sediments at different concentrations may be reported in terms of an LC50 (median lethal concentration), an EC50 (median effect concentration), an IC50 (inhibition concentration), or as a NOEC (no observed effect concentration) or LOEC (lowest observed effect concentration). In some cases, results of bioaccumulation tests may also be reported in terms of a Biota-sediment Accumulation Factor (BSAF) (Ankley et al., 1992a; Ankley et al., 1992b).

1.1.6 Evaluating effect concentrations for chemicals in sediment requires knowledge of factors controlling their bioavailability. Similar concentrations of a chemical in units of mass of chemical per mass of sediment dry weight often exhibit a range in toxicity in different sediments (Di Toro et al., 1990; Di Toro et al., 1991). Effect concentrations of chemicals in sediment have been correlated to interstitial water concentrations, and effect concentrations in interstitial water are often similar to effect concentrations in water-only exposures. The bioavailability of nonionic organic compounds in sediment is often inversely correlated with the organic carbon concentration. Whatever the route of exposure, these correlations of effect concentrations to interstitial water concentrations indicate that predicted or measured concentrations in interstitial water can be used to quantify the exposure concentration to an organism. Therefore, information on partitioning of chemicals between solid and liquid phases of sediment is useful for establishing effect concentrations (Di Toro et al., 1991).

1.1.7 Field surveys can be designed to provide either a qualitative reconnaissance of the distribution of sediment contamination or a quantitative statistical comparison of contamination among sites. Surveys of sediment toxicity or bioaccumulation are usually part of more comprehensive analyses of biological, chemical, geological, and hydrographic data. Statistical correlations may be improved and sampling costs may be reduced if subsamples are taken simultaneously for sediment tests, chemical analyses, and benthic community structure.

1.1.8 Table 1.1 lists several approaches the USEPA has considered for the assessment of sediment quality (USEPA, 1992c). These approaches include (1) equilibrium partitioning, (2) tissue residues, (3) interstitial water toxicity, (4) benthic community structure, (5) whole-sediment toxicity and sediment-spiking tests, (6) Sediment Quality Triad, and (7) sediment quality guidelines (see Chapman, 1989 and USEPA, 1989a; USEPA, 1990a; USEPA, 1990b; USEPA, 1992b for a critique of these methods). The sediment assessment approaches listed in Table 1.1 can be classified as numeric (e.g., equilibrium partitioning), descriptive (e.g., whole-sediment toxicity tests), or a combination of numeric and descriptive approaches (e.g., Effect Range Median; USEPA, 1992c). Numeric methods can be used to derive chemical-specific equilibrium partitioning sediment guidelines (ESGs) or other sediment quality guidelines (SQGs). Descriptive methods such as toxicity tests with field-collected sediment cannot be used alone to develop numerical ESGs or other SQGs for individual chemicals. Although each approach can be used to make site-specific decisions, no one single approach can adequately address sediment quality. Overall, an integration of several methods using the weight of evidence is the most desirable approach for assessing the effects of contaminants associated with sediment (Long and Morgan, 1990; MacDonald et al., 1996; Ingersoll et al., 1996; 1997). Hazard evaluations integrating data from laboratory exposures, chemical analyses, and benthic community assessments provide strong complementary evidence of the degree of pollution-induced degradation in aquatic communities (Chapman et al., 1992; Chapman et al., 1997; Burton, 1991).

1.2 Program Applicability

1.2.1 The USEPA has authority under a variety of statutes to manage contaminated sediments (Table 1.2 and USEPA, 1990e). USEPA's Contaminated Sediment Management Strategy (USEPA, 1998) establishes the following four goals for contaminated sediments and describes actions that the Agency intends to take to accomplish these goals: (1) to prevent further contamination of sediments that may cause unacceptable ecological or human health risks; (2) when practical, to clean up existing sediment contamination that adversely affects the Nation's waterbodies or their uses, or that causes other significant effects on human health or the environment; (3) to ensure that sediment dredging and the disposal of dredged material continue to be managed in an environmentally sound manner; and (4) to develop and consistently apply methodologies for analyzing contaminated sediments. The Agency plans to employ its pollution prevention and source control programs to address the first goal. To accomplish the second goal, USEPA will consider a range of risk management alternatives to reduce the volume and effects of existing contaminated sediments, including *in-situ* containment and contaminated

Table 1.1 Sediment Quality Assessment Procedures[1]

Method	Type: Numeric	Type: Descriptive	Type: Combination	Approach
Equilibrium Partitioning	*			A sediment quality value for a given contaminant is determined by calculating the sediment concentration of the contaminant that corresponds to an interstitial water concentration equivalent to the USEPA water-quality criterion for the chemical.
Tissue Residues	*			Safe sediment concentrations of specific chemicals are established by determining the sediment chemical concentration that results in acceptable tissue residues.
Interstitial Water Toxicity	*	*	*	Toxicity of interstitial water is quantified and identification evaluation procedures are applied to identify and quantify chemical components responsible for sediment toxicity.
Benthic Community Structure		*		Environmental degradation is measured by evaluating alterations in benthic community structure.
Whole-sediment Toxicity and Sediment Spiking	*	*	*	Test organisms are exposed to sediments that may contain known or unknown quantities of potentially toxic chemicals. At the end of a specified time period, the response of the test organisms is examined in relation to a specified endpoint. Dose-response relationships can be established by exposing test organisms to sediments that have been spiked with known amounts of chemicals or mixtures of chemicals.
Sediment Quality Triad	*	*	*	Sediment chemical contamination, sediment toxicity, and benthic community structure are measured on the same sediment sample. Correspondence between sediment chemistry, toxicity, and field effects is used to determine sediment concentrations that discriminate conditions of minimal, uncertain, and major biological effects.
Sediment Quality Guidelines	*	*	*	The sediment concentration of contaminanants associated with toxic responses measured in laboratory exposures or in field assessments (i.e., Apparent Effect Threshold (AET), Effect Range Median (ERM), Probable Effect Level (PEL)).

[1] Modified from USEPA (1992c)

sediment removal. Finally, the Agency is developing tools for use in pollution prevention, source control, remediation, and dredged material management to meet the collective goals. These tools include national inventories of sediment quality and environmental releases of contaminants, numerical assessment guidelines to evaluate contaminant concentrations, and standardized bioassays to evaluate the bioaccumulation and toxicity potential of sediment samples.

1.2.2 The Clean Water Act (CWA) is the single most important law dealing with environmental quality of surface waters in the United States. The objective of the CWA is to restore and maintain the chemical, physical, and biological integrity of the Nation's waters (CWA, Section 101). Federal and state monitoring programs traditionally have focused on evaluating water column problems caused by point source dischargers. Findings in the National Sediment Quality Survey, Volume I of the first biennial report to Congress on sediment quality in the U.S., indicate that this focus needs to be expanded to include sediment quality impacts (Section 1.1.2 and USEPA, 1997a).

1.2.3 The Office of Water (OW), the Office of Prevention, Pesticides, and Toxic Substances (OPPTS), the Office of Solid Waste (OSW), and the Office of Emergency and Remedial Response (OERR) are all committed to the principle of consistent tiered testing described in the Contaminated Sediment Management Strategy (USEPA, 1998). Agency-wide consistent testing is desirable because all USEPA programs will use standard methods to evaluate health risk and produce comparable data. It will also provide the basis for uniform cross-program decision-making within the USEPA. Each program will, however, retain the flexibility of deciding whether identified risks would trigger regulatory actions.

1.2.4 Tiered testing refers to a structured, hierarchical procedure for determining data needs relative to decision-making that consists of a series of tiers, or levels, of investigative intensity. Typically, increasing tiers in a tiered testing framework involve increased information and decreased uncertainty (USEPA, 1998). Each EPA program office intends to develop guidance for interpreting the tests conducted within the tiered framework and to explain how information within each tier would trigger

Table 1.2 Statutory Needs for Sediment Quality Assessment[1]

Law[2]	Area of Need
CERCLA	• Assessment of need for remedial action with contaminated sediments; assessment of degree of cleanup required, disposition of sediments
CWA	• National Pollutant Discharge Elimination System (NPDES) permitting, especially under Best Available Technology (BAT) in water-quality-limited water • Section 403(c) criteria for ocean discharges; mandatory additional requirements to protect marine environment • Section 301(g) waivers for publicly owned treatment works (POTWs) discharging to marine waters • Section 404 permits for dredge and fill activities (administered by the U.S. Army Corps of Engineers [USACE])
FIFRA	• Reviews of uses for new and existing chemicals • Pesticide labeling and registration
MPRSA	• Permits for ocean dumping
NEPA	• Preparation of environmental impact statements for projects with surface water discharges
TSCA	• Section 5: Premanufacture notification reviews for new industrial chemicals • Sections 4, 6, and 8: Reviews for existing industrial chemicals
RCRA	• Assessment of suitability (and permitting of) on-land disposal or beneficial use of contaminated sediments considered "hazardous"

[1] Modified from Dickson et al., 1987 and Southerland et al., 1992.
[2] CERCLA Comprehensive Environmental Response, Compensation and Liability Act (Superfund).
CWA Clean Water Act.
FIFRA Federal Insecticide, Fungicide, and Rodenticide Act.
MPRSA Marine Protection, Resources and Sanctuary Act.
NEPA National Environmental Policy Act.
TSCA Toxic Substances Control Act.
RCRA Resource Conservation and Recovery Act.

regulatory action. Depending on statutory and regulatory requirements, the program specific guidance will describe decisions based on a weight of evidence approach, a pass-fail approach, or comparison to a reference site. The following two approaches are currently being used by USEPA: (1) the Office of Water-U.S. Army Corps of Engineers dredged material testing framework and (2) the OPPTS ecological risk assessment tiered testing framework. USEPA-USACE (1998a) describes the dredged material testing framework and Smrchek and Zeeman (1998) summarizes the OPPTS testing framework. A tiered testing framework has not yet been chosen for Agency-wide use, but some of the components have been identified to be standardized. These components include toxicity tests, bioaccumulation tests, sediment quality guidelines, and other measurements that may have ecological significance, including benthic community structure evaluation, colonization rate, and *in situ* sediment testing within a mesocosm (USEPA, 1992a).

1.3 Scope and Application

1.3.1 A variety of standard methods have been previously developed for assessing the toxicity of chemicals in sediments using amphipods, midges, polychaetes, oligochaetes, mayflies, or cladocerans (USEPA, 1994a; USEPA, 1994b; ASTM, 1999a; ASTM, 1999b; ASTM; 1999c; ASTM, 1999d; Environment Canada, 1997a; Environment Canada, 1997b). Several endpoints are suggested in these methods to measure effects of chemicals in sediment including survival, growth, behavior, or reproduction; however, survival of test organisms in 10-d exposures is the endpoint most commonly reported. These short-term exposures which only measure effects on survival can be used to identify high levels of contamination, but might not be able to identify moderate levels of contamination in sediments (Benoit et al., 1997; Ingersoll et al., 1998; Sibley et al., 1996; Sibley et al., 1997a; Sibley et al., 1997b; Sibley et al., 1998).

1.3.2 Procedures described in Sections 11 and 12 for conducting 10-d sediment toxicity tests with the amphipod *H. azteca* (measuring survival) and the midge *C. tentans* (measuring survival and growth) were described in the first edition of the manual (USEPA, 1994a). Section 14 of this second edition of the manual now describes a method for determining potential sublethal effects of contaminants associated with sediment on *H. azteca,* including effects on reproduction based on a procedure described by Ingersoll et al. (1998). Section 15 of this second edition of the manual now describes a method for determining sublethal endpoints in sediment tests based on a life-cycle test with *C. tentans* described by Benoit et al. (1997), Sibley et al. (1996), and Sibley et al. (1997a). Procedures are primarily described for testing freshwater sediments; however, estuarine sediments (up to 15‰ salinity) can also be tested in 10-d sediment tests with *H. azteca*.

1.3.2.1 The decision to conduct 10-d or long-term toxicity tests with *H. azteca* or *C. tentans* depends on the goal of the assessment. In some instances, sufficient information may be gained by measuring sublethal endpoints in 10-d tests. In other instances, the 10-d tests could be used to screen samples for toxicity before long-term tests are conducted. While the long-term tests are needed to determine direct effects on reproduction, measurement of growth in these toxicity tests may serve as an indirect estimate of reproductive effects of chemicals associated with sediments (Section 14.4.5 and 15.4.6.2). Additional studies are ongoing to more thoroughly evaluate the relative sensitivity between lethal and sublethal endpoints measured in 10-d tests and between sublethal endpoints measured in the long-term tests. Results of these studies and additional applications of the methods described in Sections 14 and 15 will provide data that can be used to assist in determining where application of long-term tests will be most appropriate.

1.3.2.2 Use of sublethal endpoints for assessment of contaminant risk is not unique to toxicity testing with sediments. Numerous regulatory programs require the use of sublethal endpoints in the decision-making process (Pittinger and Adams, 1997) including: (1) Water Quality Criteria (and State Standards); (2) National Pollution Discharge Elimination System (NPDES) effluent monitoring (including chemical-specific limits and sublethal endpoints in toxicity tests); (3) Federal Insecticide, Rodenticide and Fungicide Act (FIFRA) and the Toxic Substances Control Act (TSCA; tiered assessment includes several sublethal endpoints with fish and aquatic invertebrates); (4) Superfund Comprehensive Environmental Response, Compensation and Liability Act (CERCLA); (5) Organization of Economic Cooperation and Development (OECD; sublethal toxicity testing with fish and invertebrates); (6) European Economic Community (EC; sublethal toxicity testing with fish and invertebrates); and (7) the Paris Commission (behavioral endpoints).

1.3.3 Guidance for conducting 28-d bioaccumulation tests with the oligochaete *Lumbriculus variegatus* is also provided in this manual (Section 13). Overlying water is renewed daily and organisms are not fed during bioaccumulation tests. Methods are also described for determining bioaccumulation kinetics of different classes of compounds during 28-d exposures with *L. variegatus.*

1.3.4 Additional research and methods development are now in progress to (1) refine sediment Toxicity Identification Evaluation (TIE) procedures (Ankley and Thomas, 1992), (2) refine sediment spiking procedures, (3) develop *in situ* toxicity tests to assess sediment toxicity and bioaccumulation under field conditions, (4) evaluate relative sensitivity of endpoints measured in toxicity tests, (5) develop methods for additional species, (6) evaluate relationships between toxicity and bioaccumulation, and (7) produce additional data on confirmation of responses in laboratory tests with natural populations of benthic organisms. This information will be described in future editions of this manual or other USEPA manuals.

1.3.4.1 This methods manual serves as a companion to the marine sediment testing method manuals (USEPA, 1994b; USEPA, 1999).

1.3.5 Procedures described in this manual are based on the following documents: ASTM (1999a), ASTM (1999b), ASTM (1999c), ASTM (1999d), Ankley et al. (1993), Phipps et al. (1993), Call et al. (1994), USEPA (1991a), USEPA (1994a), USEPA (1994b), Ingersoll et al. (1995), Ingersoll et al. (1998), Sibley et al. (1996), Sibley et al. (1997a), Sibley et al. (1997b), and Benoit et al. (1997). This manual outlines specific test methods for evaluating the toxicity of sediments in 10-d exposures with *H. azteca* and *C. tentans.* The manual also outlines general guidance on procedures for evaluating the effects of sediment contaminants in long-term exposures with *H. azteca* and *C. tentans* and bioaccumulation of contaminants in sediment with *L. variegatus.* Some issues that may be considered in interpretation of test results are the subject of continuing research, including the influence of feeding on bioavailability, nutritional requirements of the test organisms, additional performance criteria for organism health, and confirmation of responses in laboratory tests with natural benthic populations. As additional research is completed on these and other test species, the results will be incorporated into future editions of this manual. See Section 4 for additional details.

1.3.6 General procedures described in this manual might be useful for conducting tests with other aquatic organisms; however, modifications may be necessary. Altering the procedures described in this manual may alter bioavailability and produce results that are not directly comparable with results of acceptable procedures. Comparison of results obtained using modified versions of these procedures might provide useful information concerning new concepts and procedures for conducting sediment tests with aquatic organisms (e.g., *Diporeia* spp., *Tubifex tubifex, Hexagenia* spp.). If tests are conducted with procedures different from those described in this manual, additional tests are required to determine comparability of results.

1.3.6.1 Methods have been described for culturing and testing indigenous species that may be as sensitive or more sensitive than the species recommended in this manual. However, the USEPA currently allows the use of indigenous species only where state regulations require their use or prohibit importation of the recommended species. Where state regulations prohibit importation or use of the recommended test species, permission should be requested from the appropriate regulatory agency before using indigenous species.

1.3.6.2 Where states have developed culturing and testing methods for indigenous species other than those recommended in this manual, data comparing the sensitivity of the substitute species and one or more of the recommended species must be obtained with sediments or reference toxicants to ensure that the species selected are at least as sensitive and appropriate as the recommended species.

1.3.7 Selection of Test Organisms

1.3.7.1 The choice of a test organism has a major influence on the relevance, success, and interpretation of a test. Test organism selection should be based on both environmental relevance and practical concerns (DeWitt et al., 1989; Swartz, 1989). Ideally, a test organism should (1) have a toxicological database demonstrating relative sensitivity and discrimination to a range of chemicals of concern in sediment; (2) have a database for interlaboratory comparisons of procedures (e.g., round-robin studies); (3) be in contact with sediment (e.g., water column vs. benthic organism); (4) be readily available through culture or from field collection; (5) be easily maintained in the laboratory; (6) be easily identified; (7) be ecologically or economically important; (8) have a broad geographical distribution, be indigenous (either present or historical) to the site being evaluated, or have a niche similar to organisms of concern (e.g., similar feeding guild or behavior to the indigenous organisms); (9) be tolerant of a broad range of sediment physico-chemical characteristics (e.g., grain size); and (10) be compatible with selected exposure methods and endpoints (Table 1.3, ASTM, 1998d). The method should also be (11) peer reviewed (e.g., journal articles, ASTM guides) and (12) confirmed with responses with natural populations of benthic organisms (Sections 1.3.7.9 and 1.3.8.5).

1.3.7.2 Of these criteria (Table 1.3), a database demonstrating relative sensitivity to chemicals, contact with sediment, ease of culture in the laboratory, interlaboratory comparisons, tolerance to varying sediment physicochemical characteristics, and confirmation with responses of natural benthic populations were the primary criteria used for selecting *H. azteca, C. tentans,* and *L. variegatus* for the current edition of this manual. Many organisms that might be appropriate for sediment testing do not now meet these selection criteria because historically little emphasis has been placed on developing standardized testing procedures for benthic organisms. A similar database must be developed in order for other organisms to be included in future editions of this manual (e.g., mayflies [*Hexagenia* spp.], other midges [*C. riparius*], other amphipods [*Diporeia* spp.], cladocerans [*Daphnia magna, Ceriodaphnia dubia*], or mollusks).

1.3.7.3 An important consideration in the selection of specific species for test method development is the existence of information concerning relative sensitivity of the organisms both to single chemicals and complex mixtures. A number of studies have evaluated the sensitivity of *H. azteca, C. tentans* and *L. variegatus,* relative to one another, as well as other commonly tested freshwater species. For example, Ankley et al. (1991b) found *H. azteca* to be as, or slightly more, sensitive than *Ceriodaphnia dubia* to a variety of sediment elutriate and pore-water samples. In that study, *L. variegatus* were less sensitive to the samples than either the amphipod or the cladoceran. West et al. (1993) found the rank sensitivity of the three species to the lethal effects of copper in sediments could be ranked (from greatest to least): *H. azteca* > *C. tentans* > *L. variegatus*. In short-term (48 to 96 h) exposures, *L. variegatus* generally was less sensitive than *H. azteca, C. dubia,* or *Pimephales promelas* to cadmium, nickel, zinc, copper, and lead (Schubauer-Berigan et al., 1993). Of the latter three species, no one was consistently the most sensitive to all five metals.

1.3.7.3.1 In a study of Great Lakes sediment, *H. azteca, C. tentans,* and *C. riparius* were among the most sensitive and discriminatory of 24 organisms tested (Burton and Ingersoll, 1994; Burton et al., 1996a; Ingersoll et al., 1993). Kemble et al. (1994) found the rank sensitivity of four species to metal-contaminated sediments to be (from greatest to least): *H. azteca* > *C. riparius* > *Oncorhynchus mykiss* (rainbow trout) > *Daphnia magna.* The relative sensitivity of the three endpoints evaluated in the *H. azteca* test with Clark Fork River sediments was (from greatest to least): length > sexual maturation > survival.

1.3.7.3.2 In 10-d water-only and whole-sediment tests, *H. azteca* and *C. tentans* were more sensitive than *D. magna* to fluoranthene (Suedel et al., 1993).

1.3.7.3.3 Water-only tests also have been conducted for 10 d with a number of chemicals using the three species described in this manual (Phipps et al., 1995; Table 1.4). All tests were flow-through exposures using a soft natural water (Lake Superior) with measured chemical concentrations that, other than the absence of sediment, were conducted under conditions (e.g., temperature, photoperiod, feeding) similar to those being described for the standard 10-d sediment test. In general, *H. azteca* was more sensitive to copper, zinc, cadmium, nickel and lead than either *C. tentans* or *L. variegatus*. *Chironomus tentans* and *H. azteca* exhibited a similar sensitivity to several of the pesticides tested. *Lumbriculus variegatus* was not tested with several of the pesticides; however, in other studies with whole sediments contaminated by DDT and associated metabolites, and in short-term (96-h) experiments with organophosphate insecticides (diazinon, chlorpyrifos), *L. variegatus* has proven to be far less sensitive than either *H. azteca* or *C. tentans*. These results highlight two important points germane to the methods in this manual. First, neither of the two test species selected for estimating sediment toxicity *(H. azteca, C. tentans)* was consistently more sensitive to all chemicals, indicating the importance of using multiple test organisms when performing sediment assessments. Second, *L. variegatus* appears to be relatively insensitive to most of the test chemicals, which perhaps is a positive attribute for an organism used in bioaccumulation tests.

1.3.7.3.4 Using the data from Table 1.4, sensitivity of *H. azteca, C. tentans* and *L. variegatus* can be evaluated relative to other freshwater species. For this analysis, acute and chronic toxicity data from water quality criteria (WQC) documents for copper, zinc, cadmium, nickel, lead, DDT, dieldrin and chlorpyrifos, and toxicity information from the AQUIRE database (AQUIRE, 1992) for DDD and DDE, were compared to assay results for the three species (Phipps et al., 1995). The sensitivity of *H. azteca*

Table 1.3 Rating of Selection Criteria for Freshwater Sediment Toxicity Testing Organisms[1]

Criterion	*Hyalella azteca*	*Diporeia* spp.	*Chironomus tentans*	*Chironomus riparius*	*Lumbriculus variegatus*	*Tubifex tubifex*	*Hexagenia* spp.	Mollusks	*Daphnia* spp. and *Ceriodaphnia* spp.
Relative sensitivity toxicity database	+	-	+	-	+	-	-	-	-
Round-robin studies conducted	+	-	+	-	-	-	-	-	-
Contact with sediment	+	+	+	+	+	+	+	+	-
Laboratory culture	+	-	+	+	+	+	-	-	+
Taxonomic identification	+/-	+/-	+/-	+/-	+	+	+	+	+
Ecological importance	+	+	+	+	+	+	+	+	+
Geographical distribution	+	+/-	+	+	+	+	+	+	+/-
Sediment physico-chemical tolerance	+	+	+/-	+	+	+	-	+	NA
Response confirmed with benthic populations	+	+	+	+	+	+	+	-	+
Peer reviewed	+	+	+	+	+	+	+	-	+/-
Endpoints[2] monitored	S, G, M, R	S, B, A	S, G, E, R	S, G, E	B, S, R	S, R	S, G	B	S, G, R

[1] A "+" or "-" rating indicates a positive or negative attribute
[2] S = Survival, G = Growth, B = Bioaccumulation, A = Avoidance, R = Reproduction, M = Maturation, E = Emergence, NA = not applicable

Table 1.4 Water-only, 10-d LC50 (µg/L) Values for *Hyalella azteca, Chironomus tentans,* and *Lumbriculus variegatus* [1]

Chemical	*H. azteca*	*C. tentans*	*L. variegatus*
Copper	35	54	35
Zinc	73	1,125[2]	2,984
Cadmium	2.8[3]	NT[4]	158
Nickel	780	NT	12,160
Lead	<16	NT	794
p,p′-DDT	0.07	1.23	NT
p,p′-DDD	0.17	0.18	NT
p,p′-DDE	1.39	3.0	>3.3
Dieldrin	7.6	1.1	NT
Chlorpyrifos	0.086	0.07	NT

[1] Chemicals tested at ERL-Duluth in soft water—hardness 45 mg/L as $CaCO_3$ at pH 7.8 to 8.2 (Phipps et al., 1995).
[2] 50% mortality at highest concentration tested.
[3] 70% mortality at lowest concentration tested.
[4] NT = not tested.

to metals and pesticides, and *C. tentans* to pesticides was comparable to chronic toxicity data generated for other test species. This was not completely unexpected given that the 10-d exposures used for these two species are likely more similar to chronic partial life-cycle tests than the 48- to 96-h exposures traditionally defined as acute in WQC documents. Interestingly, in some instances (e.g., dieldrin, chlorpyrifos), LC50 data generated for *H. azteca* or *C. tentans* were comparable to or lower than any reported for other freshwater species in the WQC documents. This observation likely is a function not only of the test species, but of the test conditions; many of the tests on which early WQC were based were static, rather than flow-through, and utilized unmeasured contaminant concentrations.

1.3.7.4 Relative species sensitivity frequently varies among chemicals; consequently, a battery of tests including organisms representing different trophic levels may be needed to assess sediment quality (Craig, 1984;

Williams et al., 1986a; Long et al., 1990; Ingersoll et al., 1990; Burton and Ingersoll, 1994; Burton et al., 1996a; USEPA, 1989c). For example, Reish (1988) reported the relative toxicity of six metals (As, Cd, Cr, Cu, Hg, and Zn) to crustaceans, polychaetes, pelecypods, and fishes and concluded that no single species or group of test organisms was the most sensitive to all of the metals.

1.3.7.5 Measurable concentrations of ammonia are common in the pore water of many sediments and have been found to be a common cause of toxicity in pore water (Jones and Lee, 1988; Ankley et al., 1990; Schubauer-Berigan and Ankley, 1991). Acute toxicity of ammonia to *H. azteca*, *C. tentans*, and *L. variegatus* has been evaluated in several studies. As has been found for many other aquatic organisms, the toxicity of ammonia to *C. tentans* and *L. variegatus* has been shown to be dependent on pH. Four-day LC50 values for *L. variegatus* in water-column (no sediment) exposures ranged from 6.6 to 390 mg/L total ammonia as pH was increased from 6.3 to 8.6 (Schubauer-Berigan et al., 1995). For *C. tentans*, 4-d LC50 values ranged from 82 to 370 mg/L total ammonia over a similar pH range (Schubauer-Berigan et al., 1995). Ankley et al. (1995) reported that the toxicity of ammonia to *H. azteca* (also in water-only exposures) showed differing degrees of pH-dependence in different test waters. Toxicity was not pH dependent in soft reconstituted water, with 4-d LC50 values of about 20 mg/L at pH ranging from 6.5 to 8.5. In contrast, ammonia toxicity in hard reconstituted water exhibited substantial pH dependence with LC50 values decreasing from >200 to 35 mg/L total ammonia over the same pH range. Borgmann and Borgmann (1997) later showed that the variation in ammonia toxicity across these waters could be attributed to differences in sodium and potassium content, which appear to influence the toxicity of ammonia to *H. azteca*.

1.3.7.5.1 Although these studies provide benchmark concentrations that may be of concern in sediment pore waters, additional studies by Whiteman et al. (1996) indicated that the relationship between water-only LC50 values and those measured in sediment exposures differs among organisms. In sediment exposures, the 10-d LC50 for *L. variegatus* and *C. tentans* occurred when sediment pore water reached about 150% of the LC50 determined from water-only exposures. However, experiments with *H. azteca* showed that the 10-d LC50 was not reached until pore water concentrations were nearly 10 times the water-only LC50, at which time the ammonia concentration in the overlying water was equal to the water-only LC50. The authors attribute this discrepancy to avoidance of sediment by *H. azteca*. Thus, although it appears that water-only LC50 values may provide suitable screening values for potential ammonia toxicity, higher concentrations may be necessary to actually induce ammonia toxicity in sediment exposures, particularly for *H. azteca*. Further, these data underscore the importance of measuring the pH of pore water when ammonia toxicity may be of concern. Ankley and Schubauer-Berigan (1995) and Besser et al. (1998) describe procedures for conducting toxicity identification evaluations (TIEs) for pore-water or whole-sediment samples to determine whether ammonia is contributing to the toxicity of sediment samples.

1.3.7.6 Sensitivity of a species to chemicals is also dependant on the duration of the exposure and the endpoints evaluated. Sections 14.4 and 15.4 describe results of studies which demonstrate the utility of measuring sublethal endpoints in sediment toxicity tests with *H. azteca* and *C. tentans.*

1.3.7.7 The sensitivity of an organism to chemicals should be balanced with the concept of discrimination (Burton and Ingersoll, 1994; Burton et al., 1996). The response of a test organism should provide discrimination between different levels of contamination.

1.3.7.8 The sensitivity of an organism is related to the route of exposure and biochemical response to chemicals. Sediment-dwelling organisms can receive exposure from three primary sources: interstitial water, sediment particles, and overlying water. Food type, feeding rate, assimilation efficiency, and clearance rate will control the dose of chemicals from sediment. Benthic invertebrates often selectively consume different particle sizes (Harkey et al., 1994) or particles with higher organic carbon concentrations, which may have higher chemical concentrations. Grazers and other collector-gatherers that feed on aufwuchs, or surface films, and detritus may receive most of their body burden directly from materials attached to sediment or from actual sediment ingestion. In amphipods (Landrum, 1989) and clams (Boese et al., 1990), uptake through the gut can exceed uptake across the gills of certain hydrophobic compounds. Organisms in direct contact with sediment may also accumulate chemicals by direct adsorption to the body wall or by absorption through the integument (Knezovich et al., 1987).

1.3.7.9 Despite the potential complexities in estimating the dose that an animal receives from sediment, the toxicity and bioaccumulation of many chemicals in sediment such as Kepone®, fluoranthene, organochlorines, and metals have been correlated with either the concentration of these chemicals in interstitial water or, in the case of nonionic organic chemicals, in sediment on an organic-carbon normalized basis (Di Toro et al., 1990; Di Toro et al., 1991). The relative importance of whole sediment and interstitial water routes of exposure depends on the test organism and the specific chemical (Knezovich et al., 1987). Because benthic communities contain a diversity of organisms, many combinations of exposure routes can be important. Therefore, behavior and feeding habits of a test organism can influence its ability to accumulate chemicals from sediment and should be considered when selecting test organisms for sediment testing.

1.3.7.10 The response of *H. azteca* and *C. tentans* in laboratory toxicity studies has been compared with the response of natural benthic populations.

1.3.7.10.1 Chironomids were not found in sediment samples that decreased growth of *C. tentans* by 30% or

more in 10-d laboratory toxicity tests (Giesy et al., 1988). Wentsel et al. (1977a, 1977b, 1978) reported a correlation between responses of *C. tentans* in laboratory tests and the abundance of *C. tentans* in metal-contaminated sediments.

1.3.7.10.2 Canfield et al. (1994, 1996, 1998) evaluated the composition of benthic invertebrate communities in sediments for the following areas: (1) three Great Lakes Areas of Concern (AOC; Buffalo River, NY; Indiana Harbor, IN; Saginaw River, MI), (2) the upper Mississippi River, and (3) the Clark Fork River located in Montana. Results of these benthic community assessments were compared to sediment chemistry and toxicity (28-d sediment exposures with *H. azteca* which monitored effects on survival, growth, and sexual maturation). Good concordance was evident between measures of laboratory toxicity, sediment contamination, and benthic invertebrate community composition in extremely contaminated samples. However, in moderately contaminated samples, less concordance was observed between the composition of the benthic community and either laboratory toxicity test results or sediment contaminant concentration. Laboratory sediment toxicity tests better identified chemical contamination in sediments compared to many of the commonly used measures of benthic invertebrate community composition. Benthic measures may reflect other factors such as habitat alteration in addition to responding to contaminants. Canfield et al. (1994, 1996, 1998) identified the need to better evaluate noncontaminant factors (i.e., TOC, grain size, water depth, habitat alteration) in order to better interpret the response of benthic invertebrates to sediment contamination.

1.3.7.10.3 The results from laboratory sediment toxicity tests were compared to colonization of artificial substrates exposed *in situ* to Great Lakes sediment (Burton and Ingersoll, 1994; Burton et al., 1996a). Survival or growth of *H. azteca* and *C. tentans* in 10- to 28-d laboratory exposures were negatively correlated to percent chironomids and percent tolerant taxa colonizing artificial substrates in the field. Schlekat et al. (1994) reported generally good agreement between sediment tests with *H. azteca* and benthic community responses in the Anacostia River, Washington, D.C.

1.3.7.10.4 Sediment toxicity to amphipods in 10-d toxicity tests, field contamination, and field abundance of benthic amphipods were examined along a sediment contamination gradient of DDT (Swartz et al., 1994). Survival of *Eohaustorius estuarius, Rhepoxynius abronius,* and *H. azteca* in laboratory toxicity tests was positively correlated to abundance of amphipods in the field and, along with the survival of *H. azteca*, was negatively correlated to DDT concentrations. The threshold for 10-d sediment toxicity in laboratory studies was about 300 µg DDT (+metabolites)/g organic carbon. The threshold for abundance of amphipods in the field was about 100 µg DDT (+metabolites)/g organic carbon. Therefore, correlations between toxicity, contamination, and field populations indicate that short-term sediment toxicity tests can provide reliable evidence of biologically adverse sediment contamination in the field, but may be underprotective of sublethal effects.

1.3.8 Selection of Organisms for Sediment Bioaccumulation Testing

1.3.8.1 Several studies have demonstrated that hydrophobic organic compounds are bioaccumulated from sediment by freshwater infaunal organisms, including larval insects *(C. tentans,* Adams et al., 1985; Adams, 1987; *Hexagenia limbata,* Gobas et al., 1989), oligochaetes *(Tubifex tubifex* and *Limnodrilus hoffmeisteri,* Oliver, 1984; Oliver, 1987; Connell et al., 1988), and by marine organisms (polychaetes, *Nephtys incisa;* mollusks, *Mercenaria mercenaria, Yoldia limatula;* Lake et al., 1990). Consumers of these benthic organisms may bioaccumulate or biomagnify chemicals. Therefore, in addition to sediment toxicity, it may be important to examine the uptake of chemicals by aquatic organisms from contaminated sediments.

1.3.8.2 Various species of organisms have been suggested for use in studies of chemical bioaccumulation from aquatic sediments. Several criteria should be considered before a species is adopted for routine use in these types of studies (Ankley et al., 1992a; Call et al., 1994). These criteria include (1) availability of organisms throughout the year, (2) known chemical exposure history, (3) adequate tissue mass for chemical analyses, (4) ease of handling, (5) tolerance of a wide range of sediment physico-chemical characteristics (e.g., particle size), (6) low sensitivity to chemicals associated with sediment (e.g., metals, organics), (7) amenability to long-term exposures without adding food, (8) and ability to accurately reflect concentrations of chemicals in field-exposed organisms (e.g., exposure is realistic). With these criteria in mind, the advantages and disadvantages of several potential freshwater taxa for bioaccumulation testing are discussed below.

1.3.8.3 Freshwater clams provide an adequate tissue mass, are easily handled, and can be used in long-term exposures. However, few non-exotic freshwater species are available for testing. Exposure of clams is uncertain because of valve closure. Furthermore, clams are filter feeders and may accumulate lower concentrations of chemicals compared with detritivores (Lake et al., 1990). Chironomids can be readily cultured, are easy to handle, and reflect appropriate routes of exposure. However, their rapid life cycle makes it difficult to perform long-term exposures with hydrophobic compounds; also, chironomids can readily biotransform organic compounds such as benzo[a]pyrene (Harkey et al., 1994). Larval mayflies reflect appropriate routes of exposure, have adequate tissue mass for residue analysis, and can be used in long-term tests. However, mayflies cannot be continuously cultured in the laboratory and consequently are not always available for testing. Furthermore, the background concentrations of chemicals and health of field-collected individuals may be uncertain. Amphipods (e.g., *H. azteca)* can be cultured in the laboratory, are easy to handle, and reflect appropriate routes of exposure. However, their size

may be insufficient for residue analysis and *H. azteca* are sensitive to chemicals in sediment. Fish (e.g., fathead minnows) provide an adequate tissue mass, are readily available, are easy to handle, and can be used in long-term exposures. However, the route of exposure is not appropriate for evaluating the bioavailability of sediment-associated chemicals to benthic organisms.

1.3.8.4 Oligochaetes are infaunal benthic organisms that meet many of the test criteria listed above. Certain oligochaete species are easily handled and cultured, provide reasonable biomass for residue analyses, and are tolerant of varying sediment physical and chemical characteristics. Oligochaetes are exposed to chemicals via all appropriate routes of exposure, including pore water and ingestion of sediment particles. Oligochaetes need not be fed during long-term bioaccumulation exposures (Phipps et al., 1993). Various oligochaete species have been used in toxicity and bioaccumulation evaluations (Chapman et al., 1982a, Chapman et al., 1982b; Wiederholm, 1987; Kielty et al., 1988a; Kielty et al., 1988b; Phipps et al., 1993), and field populations have been used as indicators of the pollution of aquatic sediments (Brinkhurst, 1980; Spencer, 1980; Oliver, 1984; Lauritsen, 1985; Robbins et al., 1989; Ankley et al., 1992b; Brunson et al., 1993; Brunson et al., 1998). An additional desirable characteristic of *Lumbriculus variegatus* in bioaccumulation tests is that this species does not biotransform PAHs (Harkey et al., 1994).

1.3.8.5 The response of *L. variegatus* in laboratory bioaccumulation studies has been confirmed with natural populations of oligochaetes.

1.3.8.5.1 Total PCB concentrations in laboratory-exposed *L. variegatus* were similar to concentrations measured in field-collected oligochaetes from the same sites (Ankley et al., 1992b). PCB homologue patterns also were similar between laboratory-exposed and field-collected oligochaetes. The more highly chlorinated PCBs tended to have greater bioaccumulation in the field-collected organisms. In contrast, total PCBs in laboratory-exposed *(Pimephales promelas)* and field-collected *(Ictalurus melas)* fish revealed poor agreement in bioaccumulation relative to the sediment concentrations at the same sites.

1.3.8.5.2 Chemical concentrations measured in *L. variegatus* after 28-d exposures to sediment in the laboratory were compared to chemical concentrations in field-collected oligochaetes from the 13 pools of the upper Mississippi River where these sediments were collected (Brunson et al., 1998). Chemical concentrations were relatively low in sediments and tissues from the pools evaluated. Only polycyclic aromatic hydrocarbons (PAHs) and total polychlorinated biphenyls (PCBs) were frequently measured above detection limits. A positive correlation was observed between lipid-normalized concentrations of PAHs detected in laboratory-exposed *L. variegatus* and field-collected oligochaetes across all sampling locations. Rank correlations for concentrations of individual compounds between laboratory-exposed and field-collected oligochaetes were strongest for benzo(e)pyrene, perylene, benzo(b,k)-fluoranthene, and pyrene (Spearman rank correlations > 0.69). About 90% of the paired PAH concentrations in laboratory-exposed and field-collected oligochaetes were within a factor of three of one another indicating laboratory results could be extrapolated to the field with a reasonable degree of certainty.

1.4 Performance-based Criteria

1.4.1 USEPA's Environmental Monitoring Management Council (EMMC) recommended the use of performance-based methods in developing chemical analytical standards (Williams, 1993). Performance-based methods were defined by EMMC as a monitoring approach that permits the use of appropriate methods that meet pre-established demonstrated performance standards (Section 9.2).

1.4.2 The USEPA Office of Water's Office of Science and Technology and Office of Research and Development held a workshop on September 16-18, 1992 in Washington, DC to provide an opportunity for experts in the field of sediment toxicology and staff from USEPA's Regional and Headquarters program offices to discuss the development of standard freshwater and marine sediment testing procedures (USEPA, 1992a; USEPA, 1994a). Workgroup participants reached a consensus on several culturing and testing methods. In developing guidance for culturing freshwater test organisms to be included in the USEPA methods manual for sediment tests, it was agreed that no single method should be required to culture organisms. However, the consensus at the workshop was that since the success of a test depends on the health of the cultures, having healthy test organisms of known quality and age for testing was the key consideration. A performance-based criteria approach was selected as the preferred method through which individual laboratories should evaluate culture methods rather than by control-based criteria. This method was chosen to allow each laboratory to optimize culture methods and minimize effects of test organism health on the reliability and comparability of test results. See Tables 11.3, 12.3, 13.4, 14.3, and 15.3 for a listing of performance criteria for culturing and testing.

Section 2
Summary of Method

2.1 Method Description and Experimental Design

2.1.1 Method Description

2.1.1.1 This manual describes procedures for testing freshwater organisms in the laboratory to evaluate the potential toxicity or bioaccumulation of chemicals associated with whole sediments. Sediments may be collected from the field or spiked with compounds in the laboratory. Toxicity methods are outlined for two organisms, the amphipod *Hyalella azteca* and the midge *Chironomus tentans*. Methods are described for conducting 10-d toxicity tests with amphipods (Section 11) or midges (Section 12). Toxicity tests are conducted for 10 d in 300-mL chambers containing 100 mL of sediment and 175 mL of overlying water. Overlying water is added daily and test organisms are fed during the toxicity tests. The endpoints in the 10-d toxicity test with *H. azteca* and *C. tentans* are survival and growth. Procedures are primarily described for testing freshwater sediments; however, estuarine sediments (up to 15 ‰ salinity) can also be tested in 10-d toxicity tests with *H. azteca*.

2.1.1.2 Guidance is also described in the manual for conducting long-term sediment toxicity tests with *H. azteca* (Section 14) and *C. tentans* (Section 15). The long-term sediment exposures with *H. azteca* are started with 7- to 8-d-old amphipods. On Day 28, amphipods are isolated from the sediment and placed in water-only chambers where reproduction is measured on Day 35 and 42. Endpoints measured in the long-term amphipod test include survival (Day 28, 35, and 42), growth (Day 28 and 42), and reproduction (number of young per female produced from Day 28 to 42). The long-term sediment exposures with *C. tentans* start with newly hatched larvae (<24-h old) and continues through emergence, reproduction, and hatching of the F_1 generation (about 60-d exposures). Survival and growth are determined at 20 d. Starting on Day 23 to the end of the test, emergence and reproduction of *C. tentans* are monitored daily. The number of eggs per female is determined for each egg mass, which is incubated for 6 d to determine hatching success.

2.1.1.3 Guidance for conducting 28-d bioaccumulation tests with the oligochaete *Lumbriculus variegatus* is also provided in the manual. The overlying water is added daily and the test organisms are not fed during bioaccumulation tests. Section 13 also describes procedures for determining bioaccumulation kinetics of different classes of compounds during 28-d exposures with *L. variegatus*.

2.1.2 Experimental Design

The following section is a general summary of experimental design. See Section 16 for additional detail.

2.1.2.1 Control and Reference Sediment

2.1.2.1.1 Sediment tests include a control sediment (sometimes called a negative control). A control sediment is a sediment that is essentially free of contaminants, is used routinely to assess the acceptability of a test, and is not necessarily collected near the site of concern. Any contaminants in control sediment are thought to originate from the global spread of pollutants and do not reflect any substantial input from local or nonpoint sources (ASTM, 1999c). A control sediment provides a measure of test acceptability, evidence of test organism health, and a basis for interpreting data obtained from the test sediments. A reference sediment is typically collected near an area of concern (e.g., a disposal site) and is used to assess sediment conditions exclusive of material(s) of interest. Testing a reference sediment provides a site-specific basis for evaluating toxicity.

2.1.2.1.1.1 In general, the performance of test organisms in the negative control is used to judge the acceptability of a test, and either the negative control or reference sediment may be used to evaluate performance in the experimental treatments, depending on the purpose of the study. Any study in which organisms in the negative control do not meet performance criteria must be considered questionable because it suggests that adverse factors affected the test organisms. Key to avoiding this situation is using only control sediments that have a demonstrated record of performance using the same test procedure. This includes testing of new collections from sediment sources that have previously provided suitable control sediment.

2.1.2.1.1.2 Because of the uncertainties introduced by poor performance in the negative control, such studies should be repeated to insure accurate results. However, the scope or sampling associated with some studies may make it difficult or impossible to repeat a study. Some researchers have reported cases where performance in

the negative control is poor, but performance criteria are met in a reference sediment included in the study design. In these cases, it might be reasonable to infer that other samples that show good performance are probably not toxic; however, any samples showing poor performance should not be judged to have shown toxicity, since it is unknown whether the adverse factors that caused poor control performance might have also caused poor performance in the test treatments.

2.1.2.1.2 Natural geomorphological and physico-chemical characteristics such as sediment texture may influence the response of test organisms (DeWitt et al., 1988). The physico-chemical characteristics of test sediment must be within the tolerance limits of the test organism. Ideally, the limits of a test organism should be determined in advance; however, controls for factors such as grain size and organic carbon can be evaluated if the recommended limits are approached or exceeded in a test sediment. See Section 10.1 for information on physico-chemical requirements of test organisms. If the physico-chemical characteristics of a test sediment exceed the tolerance limits of the test organism, it may be desirable to include a control sediment that encompasses those characteristics. The effects of some sediment characteristics (e.g., grain size or total organic carbon) on sediment test results may be addressed with regression equations (DeWitt et al., 1988; Ankley et al., 1994a). The use of formulated sediment can also be used to evaluate physico-chemical characteristics of sediment on test organisms (Walsh et al., 1991; Suedel and Rodgers, 1994; Kemble et al., 1999; USEPA, 1998).

2.1.2.2 The experimental design depends on the purpose of the study. Variables that need to be considered include the number and type of control sediments, the number of treatments and replicates, and water-quality characteristics.

2.1.2.2.1 The purpose of the study might be to determine a specific endpoint such as an LC50 and may include a control sediment, a positive control, a solvent control, and several concentrations of sediment spiked with a chemical (see Section 8.3.2).

2.1.2.2.2 The purpose of the study might be to determine whether field-collected sediments are toxic, and may include controls, reference sediments, and test sediments. Controls are used to evaluate the acceptability of the test (Tables 11.3, 12.3, 13.4, 14.3, 15.3) and might include a control sediment, a formulated sediment (Section 7.2), a sand substrate (for *C. tentans*; Section 12.2, 15.2), or water-only exposures (for *H. azteca*; Section 14.3.7.8). Testing a reference sediment provides a site-specific basis for evaluating toxicity of the test sediments. Comparisons of test sediments to multiple reference or control sediments representative of the physical characteristics of the test sediment (i.e., grain size, organic carbon) may be useful in these evaluations. A summary of field sampling design is presented by Green (1979). See Section 16 for additional guidance on experimental design and statistics.

2.1.2.3 If the purpose of the study is to conduct a reconnaissance field survey to identify contaminated sites for further investigation, the experimental design might include only one sample from each site to allow for maximum spatial coverage. The lack of replication at a site usually precludes statistical comparisons (e.g., analysis of variance [ANOVA]) among sites, but these surveys can be used to identify contaminated sites for further study or may be evaluated using regression techniques (Sokal and Rohlf, 1981; Steel and Torrie, 1980).

2.1.2.4 In other instances, the purpose of the study might be to conduct a quantitative sediment survey of chemistry and toxicity to determine statistically significant differences between effects among control and test sediments from several sites. The number of replicates per site should be based on the need for sensitivity or power (Section 16). In a quantitative survey, replicates (separate samples from different grabs collected at the same site) would need to be taken at each site. Chemical and physical characterizations of each of these grabs would be required for each of these replicates used in sediment testing. Separate subsamples might be used to determine within-sample variability or to compare test procedures (e.g., comparative sensitivity among test organisms), but these subsamples cannot be considered to be true field replicates for statistical comparisons among sites (ASTM, 1999a).

2.1.2.5 Sediments often exhibit high spatial and temporal variability (Stemmer et al., 1990a). Therefore, replicate samples may need to be collected to determine variance in sediment characteristics. Sediments should be collected with as little disruption as possible; however, subsampling, compositing, or homogenization of sediment samples may be necessary for some experimental designs.

2.1.2.6 Site locations might be distributed along a known pollution gradient, in relation to the boundary of a disposal site, or at sites identified as being contaminated in a reconnaissance survey. Both spatial and temporal comparisons can be made. In pre-dredging studies, a sampling design can be prepared to assess the contamination of samples representative of the project area to be dredged. Such a design should include subsampling of cores taken to the project depth.

2.1.2.7 The primary focus of the physical and experimental test design, and statistical analysis of the data, is the experimental unit. The experimental unit is defined as the smallest physical entity to which treatments can be independently assigned (Steel and Torrie, 1980) and to which air and water exchange between test chambers is kept to a minimum. As the number of test chambers per treatment increases, the number of degrees of freedom and the power of a significance test increase, and therefore, the width of the confidence interval on a point estimate, such as an LC50, decreases (Section 16). Because of factors that might affect test results, all test chambers should be treated as similarly as possible. Treatments should be randomly assigned to individual test chamber

locations. Assignment of test organisms to test chambers should be impartial (Davis et al., 1998).

2.2 Types of Tests

2.2.1 Methods for conducting 10-d toxicity tests are outlined for two organisms, the amphipod *H. azteca* (Section 11) and the midge *C. tentans* (Section 12). The manual primarily describes methods for testing freshwater sediments; however, the methods described can also be used for testing *H. azteca* in estuarine sediments in 10-d tests (up to 15‰ salinity).

2.2.2 Guidance for conducting long-term toxicity tests is also outlined for *H. azteca* (Section 14) and *C. tentans* (Section 15).

2.2.3 Guidance for conducting 28-d bioaccumulation tests with the oligochaete *L. variegatus* is described in Section 13. Procedures are also described for determining bioaccumulation kinetics of different classes of compounds during 28-d exposures with *L. variegatus*.

2.3 Test Endpoints

2.3.1 Endpoints measured in the 10-d toxicity tests are survival and growth. Length or weight is reported as the average of the surviving organisms at the end of the test (Sections 11 and 12). From these data, biomass can also be calculated (dry weight of surviving organisms divided by the initial number of organisms). The rationale for evaluating biomass in toxicity testing is as follows: small differences in either growth or survival may not be statistically significantly different from the control; however, a combined estimate of biomass may increase the statistical power of the test. Although USEPA (1994c, d) describes procedures for reporting biomass as a measure of growth in effluent toxicity tests, the approach has not yet been routinely applied to sediment testing. Therefore, biomass is not listed as a primary endpoint in the methods described in Sections 11, 12, 14, and 15.

2.3.2 Endpoints measured in the long-term *H. azteca* exposures include survival (Day 28, 35, and 42), growth (Day 28 and 42), and reproduction (number of young per female produced from Day 28 to 42). The long-term sediment exposures with *C. tentans* start with newly hatched larvae (<24-h old) and continue through emergence, reproduction, and hatching of the F_1 generation (about 60-d exposures). Survival is determined at 20 d. Starting on Day 23 to the end of the test, emergence and reproduction of *C. tentans* are monitored daily. The number of eggs per female is determined for each egg mass, which is incubated for 6 d to determine hatching success.

2.3.2.1 The long-term toxicity test methods for *Hyalella azteca* and *Chironomus tentans* (Sections 14 and 15) can be used to measure effects on reproduction as well as long-term survival and growth. Reproduction is a key variable influencing the long-term sustainability of populations (Rees and Crawley, 1989) and has been shown to provide valuable and sensitive information in the assessment of sediment toxicity (Derr and Zabik, 1972; Wentsel et al., 1978; Williams et al., 1987; Postma et al., 1995; Sibley et al., 1996, 1997a; Ingersoll et al., 1998). Further, as concerns have emerged regarding the environmental significance of chemicals that can act directly or indirectly on reproductive endpoints (e.g., endocrine disrupting compounds), the need for comprehensive reproductive toxicity tests has become increasingly important. Reproductive endpoints measured in sediment toxicity tests with *H. azteca* and *C. tentans* tend to be more variable compared with those for survival or growth (Section 14.4.6 and 15.4.6). Hence, additional replicates would be required to achieve the same statistical power as for survival and growth endpoints (Section 16). The procedures described in Sections 14 and 15 include measurement of a variety of lethal and sublethal endpoints; minor modifications of the basic methods can be used in cases where only a subset of these endpoints is of interest (Sections 14.1.3 and 15.1.2).

2.3.3 Endpoints measured in bioaccumulation tests are tissue concentrations of contaminants and for some types of studies, lipid content. Behavior of test organisms should be qualitatively observed daily in all tests (e.g., avoidance of sediment).

Section 3
Definitions

3.1 Terms

The following terms were defined in Lee (1980), NRC (1989), USEPA (1989c), USEPA-USACE (1991), USEPA-USACE (1998a), ASTM (1999a), ASTM (1999b), or ASTM (1999h).

3.1.1 Technical Terms

3.1.1.1 ***Bioaccumulation.*** The net accumulation of a substance by an organism as a result of uptake from all environmental sources.

3.1.1.2 ***Bioaccumulation factor.*** Ratio of tissue residue to contaminant source concentration at steady state.

3.1.1.3 ***Bioaccumulation potential.*** Qualitative assessment of whether a contaminant is bioavailable.

3.1.1.4 ***Bioconcentration.*** The net assimilation of a substance by an aquatic organism as a result of uptake directly from aqueous solution.

3.1.1.5 ***Bioconcentration factor (BCF).*** Ratio of tissue residue to water contaminant concentration at steady state.

3.1.1.6 ***Biota-sediment accumulation factor (BSAF).*** The ratio of tissue residue to source concentration (e.g., sediment at steady state normalized to lipid and sediment organic carbon).

3.1.1.7 ***Clean.*** Denotes a sediment or water that does not contain concentrations of test materials which cause apparent stress to the test organisms or reduce their survival.

3.1.1.8 ***Concentration.*** The ratio of weight or volume of test material(s) to the weight or volume of sediment or water.

3.1.1.9 ***Contaminated sediment.*** Sediment containing chemical substances at concentrations that pose a known or suspected threat to environmental or human health.

3.1.1.10 ***Control sediment.*** A sediment that is essentially free of contaminants and is used routinely to assess the acceptability of a test. Any contaminants in control sediment may originate from the global spread of pollutants and do not reflect any substantial input from local or nonpoint sources. Comparing test sediments to control sediments is a measure of the toxicity of a test sediment beyond inevitable background contamination. Control sediment is also called a **negative control** because no toxic effects are anticipated in this treatment.

3.1.1.11 ***Depuration.*** Loss of a substance from an organism as a result of any active (e.g., metabolic breakdown) or passive process when the organism is placed into an uncontaminated environment. Contrast with Elimination.

3.1.1.12 ***Effect concentration (EC)***. The toxicant concentration that would cause an effect in a given percentage of the test population. Identical to LC when the observable adverse effect is death. For example, the EC50 is the concentration of toxicant that would cause a specified effect in 50% of the test population.

3.1.1.13 ***Elimination.*** General term for the loss of a substance from an organism that occurs by any active or passive means. The term is applicable either in a contaminated environment (e.g., occurring simultaneously with uptake) or in a clean environment. Contrast with Depuration.

3.1.1.14 ***Equilibrium partitioning sediment guidelines (ESGs).*** Numerical concentrations of chemical contaminants in sediment at or below which direct lethal or sublethal toxic effects on benthic organisms are not expected. ESGs are based on the theory that an equilibria exists among contaminant concentration in sediment pore water, contaminant associated with a binding phase in sediment, and biota. ESGs are derived by assigning a protective water-only effects concentration to the pore water (such as a Final Chronic Value), and expressing the associated equilibrium sediment concentration in terms of the principal binding phase that limits contaminant bioavailability (e.g., total organic carbon for nonionic organics or acid volatile sulfides for metals).

3.1.1.15 ***Formulated sediment.*** Mixtures of materials used to mimic the physical components of a natural sediment.

3.1.1.16 ***Inhibition concentration (IC).*** The toxicant concentration that would cause a given percent reduction in a non-quantal measurement for the test population. For

example, the IC25 is the concentration of toxicant that would cause a 25% reduction in growth for the test population, and the IC50 is the concentration of toxicant that would cause a 50% reduction.

3.1.1.17 ***Interstitial water or pore water.*** Water occupying space between sediment or soil particles.

3.1.1.18 ***k_1.*** Uptake rate coefficient from the aqueous phase, with units of g-water x g-tissue^{-1} x time^{-1}. Contrast with k_s.

3.1.1.19 ***k_2.*** Elimination rate constant, with units of time^{-1}.

3.1.1.20 ***k_s.*** Sediment uptake rate coefficient from the sediment phase, with units of g-sediment x g-tissue^{-1} x time^{-1}. Contrast with k_1.

3.1.1.21 ***K_{oc}.*** Organic carbon-water partitioning coefficient.

3.1.1.22 ***K_{ow}.*** Octanol-water partitioning coefficient.

3.1.1.23 ***Kinetic Bioaccumulation Model.*** Any model that uses uptake and/or elimination rates to predict tissue residues.

3.1.1.24 ***Lethal concentration (LC).*** The toxicant concentration that would cause death in a given percentage of the test population. Identical to EC when the observable adverse effect is death. For example, the LC50 is the concentration of toxicant that would cause death in 50% of the test population.

3.1.1.25 ***Lowest observed effect concentration (LOEC).*** The lowest concentration of a toxicant to which organisms are exposed in a test that causes an adverse effect on the test organisms (i.e., where a significant difference exists between the value for the observed response and that for the controls).

3.1.1.26 ***No observed effect concentration (NOEC).*** The highest concentration of a toxicant to which organisms are exposed in a test that causes no observable adverse effect on the test organisms (i.e., the highest concentration of a toxicant in which the value for the observed response is not statistically significantly different from the controls).

3.1.1.27 ***Overlying water.*** The water placed over sediment in a test chamber during a test.

3.1.1.28 ***Reference sediment.*** A whole sediment near an area of concern used to assess sediment conditions exclusive of material(s) of interest. The reference sediment may be used as an indicator of localized sediment conditions exclusive of the specific pollutant input of concern. Such sediment would be collected near the site of concern and would represent the background conditions resulting from any localized pollutant inputs as well as global pollutant input. This is the manner in which reference sediment is used in dredged material evaluations.

3.1.1.29 ***Reference-toxicity test.*** A test conducted with reagent-grade reference chemical to assess the sensitivity of the test organisms. Deviations outside an established normal range may indicate a change in the sensitivity of the test organism population. Reference-toxicity tests are most often performed in the absence of sediment.

3.1.1.30 ***Sediment.*** Particulate material that usually lies below water. Formulated particulate material that is intended to lie below water in a test.

3.1.1.31 ***Spiked sediment.*** A sediment to which a material has been added for experimental purposes.

3.1.1.32 ***Steady state.*** An equilibrium or "constant" tissue residue resulting from the balance of the flux of compound into and out of the organism. Operationally determined by no statistically significant difference in tissue residue concentrations from three consecutive sampling periods.

3.1.1.33 ***Whole sediment.*** Sediment and associated pore water that have had minimal manipulation. The term **bulk sediment** has been used synonymously with whole sediment.

3.1.2 Grammatical Terms

The words "must," "should," "may," "can," and "might" have very specific meanings in this manual.

3.1.2.1 "Must" is used to express an absolute requirement, that is, to state that a test ought to be designed to satisfy the specified conditions, unless the purpose of the test requires a different design. "Must" is only used in connection with the factors that directly relate to the acceptability of a test.

3.1.2.2 "Should" is used to state that the specified condition is recommended and ought to be met if possible. Although a violation of one "should" is rarely a serious matter, violation of several will often render the results questionable.

3.1.2.3 Terms such as "is desirable," "is often desirable," and "might be desirable" are used in connection with less important factors.

3.1.2.4 "May" is used to mean "is (are) allowed to," "can" is used to mean "is (are) able to," and "might" is used to mean "could possibly." Thus, the classic distinction between "may" and "can" is preserved, and "might" is never used as a synonym for either "may" or "can."

Section 4
Interferences

4.1 General Introduction

4.1.1 Interferences are characteristics of a sediment or sediment test system, aside from those related to sediment-associated chemicals of concern, that can potentially affect test organism survival, growth, or reproduction. These interferences can potentially confound interpretation of test results in two ways: (1) false-positive response, i.e., toxicity is observed in the test when contamination is not present at concentrations known to elicit a response, or there is more toxicity than expected; and (2) false-negative response, i.e., no toxicity or bioaccumulation is observed when contaminants are present at concentrations known to elicit a response, or there is less toxicity or bioaccumulation than expected.

4.1.2 There are three categories of interfering factors that can cause false-negative or false-positive responses: (1) those characteristics of sediments affecting survival independent of chemical concentration (i.e., noncontaminant factors), (2) changes in chemical bioavailability as a function of sediment manipulation or storage, and (3) the presence of indigenous organisms. Although test procedures and test organism selection criteria were developed to minimize these interferences, this section describes the nature of these interferences.

4.1.3 Because of the heterogeneity of natural sediments, extrapolation from laboratory studies to the field can sometimes be difficult (Table 4.1; Burton, 1991). Sediment collection, handling, and storage procedures may alter bioavailability and concentration of chemicals of concern by changing the physical, chemical, or biological characteristics of the sediment. Maintaining the integrity of a field-collected sediment during removal, transport, mixing, storage, and testing is extremely difficult and may complicate the interpretation of effects. Direct comparisons of organisms exposed in the laboratory and in the field would be useful to verify laboratory results. However, spiked sediment may not be representative of contaminated sediment in the field. Mixing time (Stemmer et al., 1990a), aging (Word et al., 1987; Landrum, 1989; Landrum and Faust, 1992) and the chemical form of the material can affect responses of test organisms in spiked sediment tests.

4.1.4 Laboratory testing with field-collected sediments may be useful in estimating cumulative effects and interactions of multiple chemicals in a sample. Tests with field samples usually cannot discriminate between effects of individual chemicals. Most sediment samples contain a complex matrix of inorganic and organic chemicals with many unidentified compounds. The use of Toxicity

Table 4.1 Advantages and Disadvantages for Use of Sediment Tests[1]

Advantages

- Sediment tests measure bioavailable fraction of contaminant(s).
- Sediment tests provide a direct measure of benthic effects, assuming no field adaptation or amelioration of effects.
- Limited special equipment is required for testing.
- Ten-day toxicity test methods are rapid and inexpensive.
- Legal and scientific precedence exists for use; ASTM standard guides are available.
- Sediment tests measure unique information relative to chemical analyses or benthic community analyses.
- Tests with spiked chemicals provide data on cause-effect relationships.
- Sediment toxicity tests can be applied to all chemicals of concern.
- Tests applied to field samples reflect cumulative effects of contaminants and contaminant interactions.
- Toxicity tests are amenable to confirmation with natural benthos populations.

Disadvantages

- Sediment collection, handling, and storage may alter bioavailability.
- Spiked sediment may not be representative of field contaminated sediment.
- Natural geochemical characteristics of sediment may affect the response of test organisms.
- Indigenous animals may be present in field-collected sediments.
- Route of exposure may be uncertain and data generated in sediment toxicity tests may be difficult to interpret if factors controlling the bioavailability of contaminants in sediment are unknown.
- Tests applied to field samples may not discriminate effects of individual chemicals.
- Few comparisons have been made of methods or species.
- Only a few chronic methods for measuring sublethal effects have been developed or extensively evaluated.
- Laboratory tests have inherent limitations in predicting ecological effects.

[1] Modified from Swartz (1989)

Identification Evaluations (TIE) in conjunction with sediment tests with spiked chemicals may provide evidence of causal relationships and can be applied to many chemicals of concern (Ankley and Thomas, 1992; Adams et al., 1985; USEPA, 1996b). Sediment spiking can also be used to investigate additive, antagonistic, or synergistic effects of specific chemical mixtures in a sediment sample (Swartz et al., 1988).

4.1.5 Spiked sediment may not be representative of contaminated sediment in the field. Mixing time (Stemmer et al., 1990b) and aging (Word et al., 1987; Landrum, 1989; and Landrum and Faust, 1992) of spiked sediment can affect responses of organisms.

4.1.6 Most assessments of contaminated sediment rely on short-term-lethality testing methods (e.g., ≤10 d; USEPA-USACE, 1977; USEPA-USACE, 1991; Sections 11 and 12). Short-term-lethality tests are useful in identifying "hot spots" of sediment contamination but may not be sensitive enough to evaluate moderately contaminated areas. Sediment quality assessments using sublethal responses of benthic organisms, such as effects on growth and reproduction, have been used to successfully evaluate moderately contaminated areas (Scott, 1989; Kemble et al., 1994; Ingersoll et al., 1998; Sections 14 and 15).

4.1.7 Despite the interferences discussed in this section, existing sediment test methods that include measurement of sublethal endpoints may be used to provide a rapid and direct measure of effects of contaminants on benthic communities (e.g., Canfield et al., 1996). Laboratory tests with field-collected sediment can also be used to determine temporal, horizontal, or vertical distribution of contaminants in sediment. Most tests can be completed within two to four weeks. Legal and scientific precedents exist for use of toxicity and bioaccumulation tests in regulatory decision-making (e.g., USEPA, 1986a). Furthermore, sediment tests with complex contaminant mixtures are important tools for making decisions about the extent of remedial action for contaminated aquatic sites and for evaluating the success of remediation activities.

4.2 Noncontaminant Factors

4.2.1 Results of sediment tests can be used to predict effects that may occur with aquatic organisms in the field as a result of exposure under comparable conditions. Yet motile organisms might avoid exposure in the field. Photoinduced toxicity caused by ultraviolet (UV) light may be important for some compounds associated with sediment (e.g., polycyclic aromatic hydrocarbons (PAHs); Davenport and Spacie, 1991; Ankley et al., 1994b). Fluorescent light does not contain UV light, but natural sunlight does. Lighting can therefore affect toxicological responses and is an important experimental variable for photoactivated chemicals. However, lighting typically used to conduct laboratory tests does not include the appropriate spectrum of ultraviolet radiation to photoactivate compounds (Oris and Giesy, 1985; Ankley et al., 1994b). Therefore, laboratory tests may not account for toxicity expressed by this mode of action.

4.2.2 Natural geomorphological and physico-chemical characteristics such as sediment texture may influence the response of test organisms (DeWitt et al., 1988). The physico-chemical characteristics of test sediment need to be within the tolerance limits of the test organism. Ideally, the limits of the test organism should be determined in advance; however, control samples reflecting differences in factors such as grain size and organic carbon can be evaluated if the limits are exceeded in the test sediment (Section 10.1). The effects of sediment characteristics can also be addressed with regression equations (DeWitt et al., 1988; Ankley et al., 1994a). Effects of physico-chemical characteristics of sediment on test organisms can also be evaluated by using formulated sediment for testing (Section 7.2; Walsh et al., 1991; Suedel and Rodgers, 1994; Kemble et al., 1999). See Sections 11.4, 12.4, 13.4, 14.4, and 15.4 for a discussion of the relationships between grain size of sediment and responses of test organisms.

4.2.3 A weak relationship was evident between mean reproduction of *H. azteca* in the 42-d test and grain size (Section 14.4.3; Ingersoll et al., 1998). Additional study is needed to better evaluate potential relationships between reproduction of *H. azteca* and the physical characteristics of the sediment. The weak relationship between grain size of sediment and reproduction may have been due to the fact that some of the samples with higher amounts of sand also had higher concentrations of organic chemicals compared with other samples (Ingersoll et al., 1998). *Hyalella azteca* tolerated a wide range in sediment particle size and organic matter in 10- to 28-d tests measuring effects on survival or growth (Ankley et al., 1994a; Suedel and Rodgers, 1994; Ingersoll et al., 1996; Ingersoll et al., 1998; Kemble et al., 1999; Section 14.4.3).

4.2.3.1 Until additional studies have been conducted which substantiate this lack of a correlation between physical characteristics of sediment and reproduction measured in the 42-d *H. azteca* test, it would be desirable to test control or reference sediments which are representative of the physical characteristics of field-collected sediments. Formulated sediments could be used to bracket the ranges in physical characteristics expected in the field-collected sediments being evaluated (Section 7.2). Addition of YCT should provide a minimum amount of food needed to support adequate survival, growth, and reproduction of *H. azteca* in sediments low in organic matter (Section 14.2). Without addition of food, *H. azteca* can starve during exposures (McNulty et al., 1999) making it impossible to differentiate effects of chemicals from other sediment characteristics.

4.2.4 Additional potential interferences of tests are described in Sections 11.4, 12.4, 13.4, 14.4, and 15.4.

4.3 Changes in Bioavailability

4.3.1 Sediment toxicity tests are meant to serve as an indicator of contaminant-related toxicity that might be expected under field or natural conditions. Some studies have indicated differences between results of laboratory testing and results of field testing of sediments using *in situ* exposures (Sasson-Brickson and Burton, 1991).

4.3.2 Sediment collection, handling, and storage procedures may alter contaminant bioavailability and concentration by changing the physical, chemical, or biological characteristics of the sediment. Manipulations such as mixing, homogenization, and sieving may temporarily disrupt the equilibrium of organic compounds in sediment. Similarly, oxidation of anaerobic sediments increases the availability of certain metals (Di Toro et al., 1990). Because the availability of contaminants can be a function of the degree of manipulation, this manual recommends that handling, storage, and preparation of the sediment for testing be as consistent as possible. If sieving is performed, it is done primarily to remove predatory organisms and large debris. This manipulation most likely results in a worst-case condition of heightened bioavailability yet eliminates predation as a factor that might confound test results. When sediments are sieved, it may be desirable to take samples before and after sieving (e.g., pore-water metals or DOC, AVS, TOC) to document the influence of sieving on sediment chemistry. USEPA does not recommend sieving freshwater sediments on a routine basis. See USEPA (1999) and ASTM (1999b).

4.3.3 Testing sediments at temperatures different from the field might affect contaminant solubility, partitioning coefficients, or other physical and chemical characteristics. Interaction between sediment and overlying water and the ratio of sediment to overlying water can influence bioavailability (Stemmer et al., 1990b).

4.3.4 The addition of food, water, or solvents to the test chambers might obscure the bioavailability of contaminants in sediment or might provide a substrate for bacterial or fungal growth (Harkey et al., 1997). Without addition of food, the test organisms may starve during exposures (Ankley et al., 1994a; McNulty et al., 1999). However, the addition of food may alter the availability of the contaminants in the sediment (Wiederholm et al., 1987, Harkey et al., 1994) depending on the amount of food added, its composition (e.g., TOC), and the chemical(s) of interest.

4.3.5 Depletion of aqueous and sediment-sorbed contaminants resulting from uptake by an organism or absorption to a test chamber can also influence availability. In most cases, the organism is a minor sink for contaminants relative to the sediment. However, within the burrow of an organism, sediment desorption kinetics might limit uptake rates. Within minutes to hours, a major portion of the total chemical can be inaccessible to the organisms because of depletion of available residues. The desorption of a particular compound from sediment may range from easily reversible (labile; within minutes) to irreversible (non-labile; within days or months; Karickhoff and Morris, 1985). Interparticle diffusion or advection and the quality and quantity of sediment organic carbon can also affect sorption kinetics.

4.3.6 The route of exposure may be uncertain, and data from sediment tests may be difficult to interpret if factors controlling the bioavailability of contaminants in sediment are unknown. Bulk-sediment chemical concentrations may be normalized to factors other than dry weight. For example, concentrations of nonionic organic compounds might be normalized to sediment organic-carbon content (USEPA, 1992c) and certain metals normalized to acid volatile sulfides (Di Toro et al., 1990). Even with the appropriate normalizing factors, determination of toxic effects from ingestion of sediment or from dissolved chemicals in the interstitial water can still be difficult (Lamberson and Swartz, 1988).

4.4 Presence of Indigenous Organisms

4.4.1 Indigenous organisms may be present in field-collected sediments. An abundance of the same organism or organisms taxonomically similar to the test organism in the sediment sample may make interpretation of treatment effects difficult. For example, growth of amphipods, midges, or mayflies may be reduced if high numbers of oligochaetes are in a sediment sample (Reynoldson et al., 1994). Previous investigators have inhibited the biological activity of sediment with sieving, heat, mercuric chloride, antibiotics, or gamma irradiation (see ASTM, 1999b). However, further research is needed to determine effects on contaminant bioavailability or other modifications of sediments from treatments such as those used to remove or destroy indigenous organisms.

Section 5
Health, Safety, and Waste Management

5.1 General Precautions

5.1.1 Development and maintenance of an effective health and safety program in the laboratory requires an ongoing commitment by laboratory management and includes (1) the appointment of a laboratory health and safety officer with the responsibility and authority to develop and maintain a safety program, (2) the preparation of a formal written health and safety plan, which is provided to each laboratory staff member, (3) an ongoing training program on laboratory safety, and (4) regular safety inspections.

5.1.2 This manual addresses procedures that may involve hazardous materials, operations, and equipment, but it does not purport to address all of the safety problems associated with their use. It is the responsibility of the user to establish appropriate safety and health practices, and determine the applicability of regulatory limitations before use. While some safety considerations are included in this manual, it is beyond the scope of this manual to encompass all safety requirements necessary to conduct sediment tests.

5.1.3 Collection and use of sediment may involve substantial risks to personal safety and health. Contaminants in field-collected sediment may include carcinogens, mutagens, and other potentially toxic compounds. Inasmuch as sediment testing is often begun before chemical analyses can be completed, worker contact with sediment needs to be minimized by (1) using gloves, laboratory coats, safety glasses, face shields, and respirators as appropriate, (2) manipulating sediment under a ventilated hood or in an enclosed glove box, and (3) enclosing and ventilating the exposure system. Personnel collecting sediment samples and conducting tests should take all safety precautions necessary for the prevention of bodily injury and illness that might result from ingestion or invasion of infectious agents, inhalation or absorption of corrosive or toxic substances through skin contact, and asphyxiation because of lack of oxygen or presence of noxious gases.

5.1.4 Before beginning sample collection and laboratory work, personnel should determine that all required safety equipment and materials have been obtained and are in good condition.

5.2 Safety Equipment

5.2.1 Personal Safety Gear

5.2.1.1 Personnel should use appropriate safety equipment, such as rubber aprons, laboratory coats, respirators, gloves, safety glasses, face shields, hard hats, and safety shoes.

5.2.2 Laboratory Safety Equipment

5.2.2.1 Each laboratory should be provided with safety equipment such as first aid kits, fire extinguishers, fire blankets, emergency showers, and eye wash stations.

5.2.2.2 All laboratories should be equipped with a telephone to enable personnel to summon help in case of emergency.

5.3 General Laboratory and Field Operations

5.3.1 Laboratory personnel should be trained in proper practices for handling and using chemicals that are encountered during procedures described in this manual. Routinely encountered chemicals include acids, organic solvents, and standard materials for reference-toxicity tests. Special handling and precautionary guidance in Material Safety Data Sheets should be followed for reagents and other chemicals purchased from supply houses.

5.3.2 Work with some sediment may require compliance with rules pertaining to the handling of hazardous materials. Personnel collecting samples and performing tests should not work alone.

5.3.3 It is advisable to wash exposed parts of the body with bactericidal soap and water immediately after collecting or manipulating sediment samples.

5.3.4 Strong acids and volatile organic solvents should be used in a fume hood or under an exhaust canopy over the work area.

5.3.5 An acidic solution should not be mixed with a hypochlorite solution because hazardous vapors might be produced.

5.3.6 To prepare dilute acid solutions, concentrated acid should be added to water, not vice versa. Opening a bottle of concentrated acid and adding concentrated acid to water should be performed only under a fume hood.

5.3.7 Use of ground-fault systems and leak detectors is strongly recommended to help prevent electrical shocks. Electrical equipment or extension cords not bearing the approval of Underwriter Laboratories should not be used. Ground-fault interrupters should be installed in all “wet” laboratories where electrical equipment is used.

5.3.8 All containers should be adequately labeled to identify their contents.

5.3.9 Good housekeeping contributes to safety and reliable results.

5.4 Disease Prevention

5.4.1 Personnel handling samples that are known or suspected to contain human wastes should be given the opportunity to be immunized against hepatitis B, tetanus, typhoid fever, and polio. Thorough washing of exposed skin with bactericidal soap should follow handling these samples.

5.5 Safety Manuals

5.5.1 For further guidance on safe practices when handling sediment samples and conducting toxicity tests, check with the permittee and consult general industrial safety manuals including USEPA (1986b) and Walters and Jameson (1984).

5.6 Pollution Prevention, Waste Management, and Sample Disposal

5.6.1 It is the laboratory’s responsibility to comply with the federal, state, and local regulations governing the waste management, particularly hazardous waste identification rules and land disposal restrictions, and to protect the air, water and land by minimizing and controlling all releases from fume hoods and bench operations. Also, compliance is required with any sewage discharge permits and regulations. For further information on waste management, consult “The Waste Management Manual for Laboratory Personnel” available from the American Chemical Society’s Department of Government Relations and Science Policy, 1155 16th Street N.W., Washington, D.C. 20036.

5.6.2 Guidelines for the handling and disposal of hazardous materials should be strictly followed. The federal government has published regulations for the management of hazardous waste and has given the states the option of either adopting those regulations or developing their own. If states develop their own regulations, they are required to be at least as stringent as the federal regulations. As a handler of hazardous materials, it is a laboratory's responsibility to know and comply with the applicable state regulations. Refer to The Bureau of National Affairs Inc., (1986) for the citations of the federal requirements.

5.6.3 Substitution of nonhazardous chemicals and reagents should be encouraged and investigated whenever possible. For example, use of a nonhazardous compound for a positive control in reference-toxicity tests is advisable. Reference-toxicity tests with copper can provide appropriate toxicity at concentrations below regulated levels.

Section 6
Facilities, Equipment, and Supplies

6.1 General

6.1.1 Before a sediment test is conducted in any test facility, it is desirable to conduct a "nontoxicant" test with each test species in which all test chambers contain a control sediment (sometimes called the negative control) and clean overlying water. Survival, growth, or reproduction of the test organisms will demonstrate whether facilities, water, control sediment, and handling techniques are adequate to result in acceptable species-specific control numbers. Evaluations may also be made on the magnitude of between-chamber variance in a test. See Section 9.14.

6.2 Facilities

6.2.1 The facility must include separate areas for culturing test organisms and sediment testing to reduce the possibility of contamination by test materials and other substances, especially volatile compounds. Holding, acclimation, and culture chambers should not be in a room where sediment tests are conducted, stock solutions or sediments are prepared, or equipment is cleaned. Test chambers may be placed in a temperature-controlled recirculating water bath, environmental chamber, or equivalent facility with temperature control. An enclosed test system is desirable to provide ventilation during tests to limit exposure of laboratory personnel to volatile substances.

6.2.2 Light of the quality and luminance normally obtained in the laboratory is adequate (about 100 to 1000 lux using wide-spectrum fluorescent lights; e.g., cool-white or daylight) has been used successfully to culture and test organisms. Lux is the unit selected for reporting luminance in this manual. Multiply units of lux by 0.093 to convert to units of foot candles. Multiply units of lux by 6.91×10^{-3} to convert to units of $\mu E/m^2/s$ (assuming an average wavelength of 550 nm ($\mu mol^{-2} s^{-1} = W m \times \lambda[nm] \times 8.36 \times 10^{-3}$); ASTM, 1999g). Luminance should be measured at the surface of the water in test chambers. A uniform photoperiod of 16L:8D can be achieved in the laboratory or in an environmental chamber using automatic timers.

6.2.3 During phases of rearing, holding, and testing, test organisms should be shielded from external disturbances such as rapidly changing light or pedestrian traffic.

6.2.4 The test facility should be well ventilated and free of fumes. Laboratory ventilation systems should be checked to ensure that return air from chemistry laboratories or sample handling areas is not circulated to culture or testing rooms, or that air from testing rooms does not contaminate culture rooms. Air pressure differentials between rooms should not result in a net flow of potentially contaminated air to sensitive areas through open or loose-fitting doors. Air used for aeration must be free of oil and fumes. Oil-free air pumps should be used where possible. Filters to remove oil, water, and bacteria are desirable. Particles can be removed from the air using filters such as BALSTON® Grade BX (Balston, Inc., Lexington, MA) or equivalent, and oil and other organic vapors can be removed using activated carbon filters (e.g., BALSTON® C-1 filter), or equivalent.

6.3 Equipment and Supplies

6.3.1 Equipment and supplies that contact stock solutions, sediment, or overlying water should not contain substances that can be leached or dissolved in amounts that adversely affect the test organisms. In addition, equipment and supplies that contact sediment or water should be chosen to minimize sorption of test materials from water. Glass, type 316 stainless steel, nylon, high-density polyethylene, polypropylene, polycarbonate, and fluorocarbon plastics should be used whenever possible to minimize leaching, dissolution, and sorption. Concrete and high-density plastic containers may be used for holding and culture chambers, and in the water-supply system. These materials should be washed in detergent, acid rinsed, and soaked in flowing water for a week or more before use. Cast-iron pipe should not be used in water-supply systems because colloidal iron will be added to the overlying water and strainers will be needed to remove rust particles. Copper, brass, lead, galvanized metal, and natural rubber must not contact overlying water or stock solutions before or during a test. Items made of neoprene rubber and other materials not mentioned above should not be used unless it has been shown that their use will not adversely affect survival, growth, or reproduction of the test organisms.

6.3.2 New lots of plastic products should be tested for toxicity by exposing organisms to them under ordinary test conditions before general use.

6.3.3 General Equipment

6.3.3.1 Environmental chamber or equivalent facility with photoperiod and temperature control (20°C to 25°C).

6.3.3.2 Water purification system capable of producing at least 1 mega-ohm water (USEPA, 1991a).

6.3.3.3 Analytical balance capable of accurately weighing to 0.01 mg.

6.3.3.4 Reference weights, Class S—for documenting the performance of the analytical balance(s). The balance(s) should be checked with reference weights that are at the upper and lower ends of the range of the weighings made when the balance is used. A balance should be checked at the beginning of each series of weighings, periodically (such as every tenth weight) during a long series of weighings, and after taking the last weight of a series.

6.3.3.5 Volumetric flasks and graduated cylinders—Class A, borosilicate glass or nontoxic plastic labware, 10 to 1000 mL for making test solutions.

6.3.3.6 Volumetric pipets—Class A, 1 to 100 mL.

6.3.3.7 Serological pipets—1 to 10 mL, graduated.

6.3.3.8 Pipet bulbs and fillers—PROPIPET® or equivalent.

6.3.3.9 Droppers, and glass tubing with fire polished edges, 4- to 6-mm ID—for transferring test organisms.

6.3.3.10 Wash bottles—for rinsing small glassware, instrument electrodes and probes.

6.3.3.11 Glass or electronic thermometers—for measuring water temperature.

6.3.3.12 National Bureau of Standards Certified thermometer (see USEPA Method 170.1; USEPA, 1979b).

6.3.3.13 Dissolved oxygen (DO), pH/selective ion, and specific conductivity meters and probes for routine physical and chemical measurements are needed. Unless a test is being conducted to specifically measure the effect of DO or conductivity, a portable field-grade instrument is acceptable.

6.3.3.14 See Table 6.1 for a list of additional equipment and supplies. Appendix C outlines additional equipment and supplies needed for conducting the long-term exposures with *C. tentans.*

6.3.4 Water-delivery System

6.3.4.1 The water-delivery system used in water-renewal testing can be one of several designs (Appendix A). The system should be capable of delivering water to each replicate test chamber. Mount and Brungs (1967) diluters have been successfully modified for sediment testing. Other diluter systems have also been useful (Ingersoll and Nelson, 1990; Maki, 1977; Benoit et al., 1993; Zumwalt et al., 1994; Brunson et al., 1998). The water-delivery system should be calibrated before the test by determining the flow rate of the overlying water. The general operation of the system should be visually checked daily throughout the length of the test. If necessary, the water-delivery system should be adjusted during the test. At any particular time during the test, flow rates through any two test chambers should not differ by more than 10%.

6.3.4.2 The overlying water can be replaced manually (e.g., siphoning); however, manual systems take more time to maintain during a test. In addition, automated systems generally result in less suspension of sediment compared to manual renewal.

6.3.5 Test Chambers

6.3.5.1 Test chambers may be constructed in several ways and of various materials, depending on the experimental design and the contaminants of interest. Clear silicone adhesives, suitable for aquaria, sorb some organic compounds that might be difficult to remove. Therefore, as little adhesive as possible should be in contact with the test material. Extra beads of adhesive should be on the outside of the test chambers rather than on the inside. To leach potentially toxic compounds from the adhesive, all new test chambers constructed using silicone adhesives should be held at least 48 h in overlying water before use in a test.

6.3.5.2 Test chambers for specific tests are described in Sections 11, 12, 13, 14, and 15.

6.3.6 Cleaning

6.3.6.1 All nondisposable sample containers, test chambers, and other equipment that have come in contact with sediment should be washed after use in the manner described below to remove surface contaminants.

1. Soak 15 min in tap water and scrub with detergent, or clean in an automatic dishwasher.
2. Rinse twice with tap water.
3. Carefully rinse once with fresh, dilute (10%, V:V) hydrochloric or nitric acid to remove scale, metals, and bases. To prepare a 10% solution of acid, add 10 mL of concentrated acid to 90 mL of deionized water.
4. Rinse twice with deionized water.
5. Rinse once with full-strength, pesticide-grade acetone to remove organic compounds (use a fume hood or canopy). Hexane might also be used as a solvent for removing nonionic organic compounds. However, acetone is preferable if only one organic solvent is used to clean equipment.
6. Rinse three times with deionized water.

Table 6.1 Equipment and Supplies for Culturing and Testing Specific Test Organisms[1]

A. Biological Supplies

Brood stock of test organisms
Active dry yeast (HA)
Cerophyl® (dried cereal leaves; HA)
Trout food pellets (HA)
Tetrafin® or Tetramin® goldfish food (CT)
Trout starter (LV)
Helisoma sp. snails (optional; LV)
Algae (e.g., *Selenastrum capricornutum, Chlorella;* CT)
Diatoms (e.g., *Navicula* sp; HA)

B. Glassware

Culture chambers
Test chambers (300-mL high-form lipless beaker; HA and CT)
Test chambers (15.8- x 29.3- x 11.7-cm, W x L x H; LV)
Juvenile holding beakers (e.g., 1 L; HA)
Crystallizing dishes or beakers (200 to 300 mL; CT)
Erlenmeyer flasks (250 and 500 mL; CT)
Larval rearing chambers (e.g., l9 L capacity; CT)
1/4" glass tubing (for aspirating flask; CT)
Glass bowls (20-cm diameter; LV)
Glass vials (10 mL; LV)
Wide-bore pipets (4- to 6-mm ID)
Glass disposable pipets
Burettes (for hardness and alkalinity determinations)
Graduated cylinders (assorted sizes, 10 mL to 2 L)

C. Instruments and Equipment

Dissecting microscope
Stainless-steel sieves (e.g., U.S. Standard No. 25, 30 35, 40, 50 mesh)
Delivery system for overlying water (See Appendix B for a listing of equipment needed for water delivery systems)
Photoperiod timers
Light meter
Temperature controllers
Thermometer
Continuous recording thermometers
Dissolved oxygen meter
pH meter
Ion-specific meter
Ammonia electrode (or ammonia test kit)
Specific-conductance meter
Drying oven
Desiccator
Balance (0.01 mg sensitivity)

C. Instruments and Equipment

Blender
Refrigerator
Freezer
Light box
Hemacytometer (HA)
Paper shredder, cutter, or scissors (CT, LV)
Tissue homogenizer (LV)
Electric drill with stainless steel auger (diameter 7.6 cm, overall length 38 cm, auger bit length 25.4 cm (Section 8.3)

D. Miscellaneous

Ventilation system for test chambers
Air supply and airstones (oil free and regulated)
Cotton surgical gauze or cheese cloth (HA)
Stainless-steel screen (no. 60 mesh, for test chambers)
Glass hole-cutting bits
Silicon adhesive caulking
Plastic mesh (110-µm mesh opening; Nytex® 110; HA)
Aluminum weighing pans (Sigma Chemical Co., St. Louis, MO)
Fluorescent light bulbs
Nalgene bottles (500 mL and 1000 mL for food preparation and storage)
Deionized water
Air line tubing
White plastic dish pan
"Coiled-web material" (3-M, St. Paul, MN; HA)
White paper toweling (for substrate; CT)
Brown paper toweling (for substrate; LV)
Screening material (e.g., Nitex® (110 mesh), window screen, or panty hose; CT)
Water squirt bottle
Dissecting probes (LV)
Dental picks (LV)
Shallow pans (plastic (light-colored), glass, stainless steel)

E. Chemicals

Detergent (nonphosphate)
Acetone (reagent grade)
Hexane (reagent grade)
Hydrochloric acid (reagent grade)
Chloroform and methanol (LV)
Copper Sulfate, Potassium Chloride
Reagents for reconstituting water
Formalin (or Notox®)
Sucrose

HA = *Hyalella azteca*
CT = *Chironomus tentans*
LV = *Lumbriculus variegatus*
[1] Appendix C outlines additional equipment and supplies for the long-term exposures with *C. tentans.*

6.3.6.2 All test chambers and equipment should be thoroughly rinsed or soaked with the dilution water immediately before use in a test.

6.3.6.3 Many organic solvents (e.g., methylene chloride) leave a film that is insoluble in water. A dichromate-sulfuric acid cleaning solution can be used in place of both the organic solvent and the acid (see ASTM, 1999e), but the solution might attack silicone adhesive and leave chromium residues on glass. An alternative to use of dichromate-sulfuric acid could be to heat glassware for 8 h at 450°C.

Section 7
Water, Formulated Sediment, Reagents, and Standards

7.1 Water

7.1.1 Requirements

7.1.1.1 Water used to test and culture organisms should be uniform in quality. Acceptable water should allow satisfactory survival, growth, or reproduction of the test organisms. Test organisms should not show signs of disease or apparent stress (e.g., discoloration, unusual behavior). If problems are observed in the culturing or testing of organisms, it is desirable to evaluate the characteristics of the water. See USEPA (1991a) and ASTM (1999a) for a recommended list of chemical analyses of the water supply.

7.1.2 Source

7.1.2.1 A natural water is considered to be of uniform quality if monthly ranges of the hardness, alkalinity, and specific conductance are less than 10% of their respective averages and if the monthly range of pH is less than 0.4. Natural waters should be obtained from an uncontaminated well or spring, if possible, or from a surface-water source. If surface water is used, the intake should be positioned to (1) minimize fluctuations in quality and contamination, (2) maximize the concentration of dissolved oxygen, and (3) ensure low concentrations of sulfide and iron. Municipal water supplies may be variable and may contain unacceptably high concentrations of materials such as copper, lead, zinc, fluoride, chlorine, or chloramines. Chlorinated water should not be used for culturing or testing because residual chlorine and chlorine-produced oxidants are toxic to many aquatic organisms. Use of tap water is discouraged unless it is dechlorinated and passed through a deionizer and carbon filter (USEPA, 1991a).

7.1.2.2 For site-specific investigations, it is desirable to have the water-quality characteristics of the overlying water as similar as possible to the site water. For certain applications the experimental design might require use of water from the site where sediment is collected.

7.1.2.3 Water that might be contaminated with facultative pathogens may be passed through a properly maintained ultraviolet sterilizer equipped with an intensity meter and flow controls or passed through a filter with a pore size of 0.45 µm or less.

7.1.2.4 Water might need aeration using air stones, surface aerators, or column aerators. Adequate aeration will stabilize pH, bring concentrations of dissolved oxygen and other gases into equilibrium with air, and minimize oxygen demand and concentrations of volatiles. Excessive aeration may reduce hardness and alkalinity of hard water (e.g., 280 mg/L hardness as $CaCO_3$; E.L. Brunson, USGS, Columbia, MO, personal communication). The concentration of dissolved oxygen in source water should be between 90 to 100% saturation to help ensure that dissolved oxygen concentrations are acceptable in test chambers.

7.1.3 Reconstituted Water

7.1.3.1 Ideally, reconstituted water should be prepared by adding specified amounts of reagent-grade chemicals to high-purity distilled or deionized water (ASTM, 1999e; USEPA, 1991a). Problems have been observed with use of reconstituted water in long-term exposures with *H. azteca* (Section 7.1.3.4.3). In some applications, acceptable high-purity water can be prepared using deionization, distillation, or reverse-osmosis units (Section 6.3.3.2; USEPA, 1991a). In some applications, test water can be prepared by diluting natural water with deionized water (Kemble et al., 1994) or by adding salts to relatively dilute natural waters.

7.1.3.2 Deionized water should be obtained from a system capable of producing at least 1 mega-ohm water. If large quantities of high quality deionized water are needed, it may be advisable to supply the laboratory grade water deionizer with preconditioned water from a mixed-bed water treatment system. Some investigators have observed that holding reconstituted water prepared from deionized water for several days before use in sediment tests may improve performance of test organisms.

7.1.3.3 Conductivity, pH, hardness, dissolved oxygen, and alkalinity should be measured on each batch of reconstituted water. The reconstituted water should be aerated before use to adjust pH and dissolved oxygen to the acceptable ranges (e.g., Section 7.1.3.4.1). USEPA (1991a) recommends using a batch of reconstituted water for two weeks.

7.1.3.4 Reconstituted Fresh Water (Smith et al., 1997)

7.1.3.4.1 To prepare 100 L of reconstituted fresh water, use the reagent-grade chemicals as follows:

1. Place about 75 L of deionized water in a properly cleaned container.

2. Add 5 g of $CaSO_4$ and 5 g of $CaCl_2$ to a 2-L aliquot of deionized water and mix (e.g., on a stir plate) for 30 min or until the salts dissolve.

3. Add 3 g of $MgSO_4$, 9.6 g $NaHCO_3$, and 0.4 g KCl to a second 2-L aliquot of deionized water and mix on a stir plate for 30 min.

4. Pour the two 2-L aliquots containing the dissolved salts into the 75 L of deionized water and fill the carboy to 100 L with deionized water.

5. Aerate the mixture for at least 24 h before use.

6. The water quality of the reconstituted water should be approximately the following: hardness, 90 to 100 mg/L as $CaCO_3$, alkalinity 50 to 70 mg/L as $CaCO_3$, conductivity 330 to 360 mS/cm, and pH 7.8 to 8.2.

7.1.3.4.2 This reconstituted fresh water (reformulated moderately hard reconstituted water) described by Smith et al. (1997) and described in the first edition of this manual (USEPA, 1994a) has been used successfully in 10-d round-robin testing with *H. azteca, C. tentans,* and *C. riparius* (Section 17). This reconstituted water has a higher proportion of chloride to sulfate compared to the reconstituted waters described in ASTM (1999e) and USEPA (1991a).

7.1.3.4.3 McNulty et al. (1999) and Kemble et al. (1998, 1999) observed poor survival of *H. azteca* in tests conducted 14 to 28 d using a variety of reconstituted waters including the reconstituted water described by Smith et al. (1997). Borgmann (1996) described a reconstituted water that was used successfully to maintain *H. azteca* in culture; however, some laboratories have not had success with reproduction of the *H. azteca* when using this reconstituted water in the 42-d test (T.J. Norberg-King, USEPA, Duluth, MN, personal communication). Research is ongoing to develop additional types of reconstituted waters suitable for *H. azteca*. Until an acceptable reconstituted water has been developed for long-term exposures with *H. azteca,* a natural water demonstrated to support adequate survival, growth, and reproduction of amphipods is recommended for use in long-term *H. azteca* exposures (Section 14.2; Ingersoll et al., 1998; Kemble et al., 1998, 1999).

7.1.3.5 Synthetic Seawater

7.1.3.5.1 Reconstituted salt water can be prepared by adding commercial sea salts, such as FORTY FATHOMS®, HW MARINEMIX®, INSTANT OCEAN®, or equivalent to deionized water.

7.1.3.5.2 A synthetic seawater formulation called GP2 is prepared with reagent grade chemicals that can be diluted with deionized water to the desired salinity (USEPA, 1994d).

7.1.3.5.3 Ingersoll et al. (1992) describe procedures for culturing *H. azteca* at salinities up to 15 ‰. Reconstituted salt water was prepared by adding INSTANT OCEAN® salts to a 25:75 (v/v) mixture of freshwater (hardness 283 mg/L as $CaCO_3$) and deionized water that was held at least two weeks before use. Synthetic seawater was conditioned by adding 6.2 mL of Frit-zyme® #9 nitrifying bacteria (*Nitromonas* sp. and *Nitrobacter* sp.; Fritz Chemical Company, Dallas, TX) to each liter of water. The cultures were maintained by using renewal procedures; 25% of the culture water was replaced weekly. *Hyalella azteca* have been used to evaluate the toxicity of estuarine sediments up to 15 ‰ salinity in 10-d exposures (Nebeker and Miller, 1988; Roach et al., 1992; Winger et al., 1993; Ingersoll et al., 1996).

7.2 Formulated Sediment

7.2.1 General Requirements

7.2.1.1 Formulated sediments are mixtures of materials that mimic the physical components of natural sediments. Formulated sediments have not been routinely applied to evaluate sediment contamination. A primary use of formulated sediment could be as a control sediment. Formulated sediments allow for standardization of sediment testing or provide a basis for conducting sediment research. Formulated sediment provides a basis by which any testing program can assess the acceptability of their procedures and facilities. In addition, formulated sediment provides a consistent measure evaluating performance-based criteria necessary for test acceptability. The use of formulated sediment eliminates interferences caused by the presence of indigenous organisms. For toxicity tests with sediments spiked with specific chemicals, the use of a formulated sediment eliminates or controls the variation in sediment physico-chemical characteristics and provides a consistent method for evaluating the fate of chemicals in sediment. See USEPA (1999) and ASTM (1999b) for additional detail regarding uses of formulated sediment.

7.2.1.2 A formulated sediment should (1) support the survival, growth, or reproduction of a variety of benthic invertebrates, (2) provide consistent acceptable biological endpoints for a variety of species, and (3) be composed of materials that have consistent characteristics. Consistent material characteristics include (1) consistency of materials from batch to batch, (2) contaminant concentrations below concentrations of concern, and (3) availability to all individuals and facilities (Kemble et al., 1999).

7.2.1.3 Physico-chemical characteristics that might be considered when evaluating the appropriateness of a formulated sediment include percent sand, percent clay, percent silt, organic carbon content, cation exchange

capacity (CEC), oxidation reduction potential (redox), pH, and carbon:nitrogen:phosphorus ratios.

7.2.2 Sources of Materials

7.2.2.1 A variety of methods describe procedures for making formulated sediments. These procedures often use similar constituents; however, they often include either a component or a formulation step that would result in variation from test facility to test facility. In addition, most of the procedures have not been subjected to standardization and consensus approval or round-robin (ring) testing. The procedure outlined by Kemble et al. (1999) below was evaluated in round-robin testing with *Hyalella azteca* and *Chironomus tentans* (Section 17.6).

7.2.2.2 Most formulated sediments include sand and clay/silt that meet certain specifications; however, some may be quite different. For example, three sources of clay and silt include Attagel® 50, ASP® 400, and ASP® 400P. Table 7.1 summarizes the characteristics of these materials. The percentage of clay ranges from 56.5 to 88.5 and silt ranges from 11.5 to 43.5. These characteristics should be evaluated when considering the materials to use in a formulated sediment.

Table 7.1 Characteristics of Three Sources of Clays and Silts Used in Formulated Sediments

Characteristic	Attagel® 50	ASP® 400	ASP® 400P
% Sand	0.0	0.01	0.0
% Clay	88.50	68.49	56.50
% Silt	11.50	31.50	43.50
Soil class	Clay	Clay	Silty clay

Note: Table 7.3 lists suppliers for these materials.

7.2.2.3 A critical component of formulated sediment is the source of organic carbon. Many procedures have used peat as the source of organic carbon. Other sources of organic carbon listed in Table 7.2 have been evaluated including humus, potting soil, maple leaves, composted cow manure, rabbit chow, cereal leaves, chlorella, trout chow, Tetramin®, Tetrafin®, and alpha cellulose. Only peat, humus, potting soil, composted cow manure, and alpha cellulose have been used successfully without fouling the overlying water in sediment testing (Kemble et al., 1999). The other sources of organic carbon listed in Table 7.2 caused dissolved oxygen concentrations to fall to unacceptable levels (Kemble et al., 1999). Kemble et al. (1999) reported that conditioning of formulated sediment was not necessary when alpha cellulose was used as a source of organic carbon to prepare sediment for use as a negative control. In addition, alpha cellulose is a consistent source of organic carbon that is relatively biologically inactive and low in concentrations of chemicals of concern. It is one of three forms of cellulose (alpha, beta, and gamma) that differ in their degree of polymerization. Alpha cellulose has the highest degree of polymerization and is the chief constituent of paper pulp. The beta and gamma forms have a much lower degree of polymerization and are known as hemicellulose. Hence, compared with other sources of organic carbon, alpha cellulose would not serve as a food source, but would serve as an organic carbon constituent for sediment to add texture or to provide a partitioning compartment for chemicals. Using alpha cellulose as a source of organic carbon for sediment-spiking studies has not been adequately evaluated. Recent work conducted by J. Besser (USGS, Columbia, MO, unpublished data) indicated that using alpha cellulose as a source or organic carbon in 21-d studies resulted in some generation of sulfide in the pore water, which may affect the bioavailability of metals spiked in sediment.

Table 7.2. Carbon, Nitrogen, Phosphorus Levels for Various Sources of Organic Carbon (Kemble et al., 1998a)

Organic carbon Source	Carbon (%)	Nitrogen (mg/g)	Phosphorus (µg/g)
Peat	47	4	0.4
Maple leaves 1	42	6	1.3
Maple leaves 2	47	3	1.7
Cow manure	30	11	8.2
Rabbit chow	40	18	0.2
Humic acid	40	3	ND[1]
Cereal leaves	47	4	0.4
Chlorella	40	41	5.7
Trout chow	43	36	11.0
Tetramin®	37	45	9.6
Tetrafin®	36	29	8.6
Alpha cellulose	30	0.7	ND

[1] Not detected.

7.2.2.4 An important consideration in the selection of an organic carbon source may be the ratio of carbon to nitrogen to phosphorus. As demonstrated in Table 7.2, percentage carbon ranged from 30 to 47, nitrogen ranged from 0.7 to 45 mg/g, and phosphorus ranged from below detection to 11 µg/g for several different carbon sources. These characteristics should be evaluated when considering the materials to use in a formulated sediment.

7.2.3 Procedure

7.2.3.1 A summary of various procedures that have been used to formulate sediment are listed below. Suppliers of various components are listed in Table 7.3.

1. Walsh et al. (1981): (1) Wash sand (Mystic White No. 85, 45, and 18—New England Silica Inc.; Note: Mystic White sands are no longer available. Kemble et al. (1999) found White Quartz sand to be an acceptable substitute; Table 7.3) and sieve into three grain sizes: coarse (500 to 1500 mm); medium (250 to 499 mm); and fine (63 to 249 mm). (2) Obtain clay and silt from Engelhard Corp. (3) Mill and sieve peat moss through an 840-mm screen. (4) Mix constituents dry in the following quantities: coarse sand (0.6%); medium sand (8.7%); fine sand (69.2%); silt (10.2%); clay (6.4%); and organic matter (4.9%).

Table 7.3 Sources of Components Used in Formulated Sediments

Component	Sources
Sand	• White Quartz sand #1 dry, #2, #3—New England Silica, Inc., South Windsor, CT (Note: Mystic White sands are no longer available. Kemble et al. (1999) found White Quartz sand to be an acceptable substitute).
	• Product No. 33094, BDH Chemical, Ltd., Poole, England
Kaolinite	• ASP 400, ASP 400P, ASP 600, ASP 900—Englehard Corporation, Edison, NJ
	• Product No. 33059, BDH Chemical, Ltd., Poole, England
Montmorillonite	• W.D. Johns, Source Clays, University of Missouri, Columbia, MO
Clay	• Lewiscraft Sculptor's Clay, available in hobby and artist supply stores
Humus	• Sims Bark Co., Inc., Tuscumbia, AL
Alpha cellulose	• Sigma Co., St. Louis, MO
Peat	• D.L. Browning Co., Mather, WI
	• Joseph Bentley, Ltd., Barrow-on-Humber, South Humberside, England
	• Mellinger's, North Lima, OH
Potting soil	• Zehr's No Name Potting Soil, Mississauga, Ontario
Humic acid	• Aldrich Chemical Co, Milwaukee, WI
Cow manure	• A.H. Hoffman, Inc., Landisville, PA
Dolomite	• Ward's Natural Science Establishment, Inc., Rochester, NY

2. Harrahy and Clements (1997): (1) Rinse peat moss then soak for 5 d in deionized water renewing water daily. (2) After acclimation for 5 d, remove all water and spread out to dry. (3) Grind moss and sieve using the following sieve sizes: 1.18 mm (discard these particles); 1.00 mm (average size 1.09 mm); 0.85 mm (average size 0.925); 0.60 (average size 0.725); 0.425 mm (average size 0.5125 mm); retainer (average size 0.2125 mm). (4) Use a mixture of sizes that provides an average particle size of 840 mm. (5) Wash medium quartz sand and dry. (6) Obtain clay and silt using ASP 400 (Englehard Corp). (7) Mix constituents dry in the following quantities: sand (850 g); silt and clay (150 g); dolomite (0.5 g); sphagnum moss (22 g); and humic acid (0.1g). (8) Mix sediment for an hour on a rolling mill and store dry until ready for use.

3. Hanes et al. (1991): (1) Sieve sand and retain two particle sizes (90 to 180 um and 180 to 250 um) which are mixed in a ratio of 2:1. (2) Dry potting soil for 24 h at room temperature and sieve through a 1-mm screen. Clay is commercially available sculptors clay. (3) Determine percent moisture of clay and soil after drying for 24 h at 60 to 100°C (correct for percent moisture when mixing materials). (4) Mix constituents by weight in the following ratios: sand mixture (42%); clay (42%); and soil (16%). (5) Autoclave after mixing in a foil-covered container for 20 min. Mixture can be stored indefinitely if kept covered after autoclaving.

4. Naylor (1993): (1) Sieve acid-washed sand to obtain a 40- to 100-mm size. (2) Obtain clay as kaolin light. (3) Grind and sieve peat moss using a 2-mm screen (peat moss which is allowed to dry out will not rehydrate and will float on the water surface). (4) Adjust for the use of moist peat moss by determining moisture content (dry 5 samples of peat at 60°C until constant weight is achieved). (5) Mix constituents by weight in the following percentages: sand (69%); kaolin (20%); peat (10% [adjust for moisture content]); and $CaCO_3$ (1%). (6) Mix for 2 h in a soil shaker and store in sealed containers.

5. Suedel and Rodgers (1994): (1) Sieve sand (Mystic White #18 and 90; Note: Mystic White sands are no longer available. Kemble et al. (1999) found White Quartz sand to be an acceptable substitute; Table 7.3) to provide three different size fractions: coarse (2.0 to 0.5 mm), medium (0.5 to 0.25 mm) and fine (0.25 to 0.05 mm). (2) Ash silt (ASP 400), clay (ASP 600 and 900), montmorillonite clay, and dolomite at 550°C for 1 h to remove organic matter. (3) Dry humus (70°C) and mill to 2.0 mm. (4) Add dolomite as 1% of the silt requirement. (5) Age materials for 7 d in flowing water before mixing. (6) Mix constituents to mimic the desired characteristics of the sediment of concern.

6. Kemble et al. (1999) describe procedures for making a variety of formulated sediments ranging in grain size and organic carbon. A sediment with 19% sand and 2% organic carbon was produced by combining: (1) 219 grams of sand (White Quartz #1 dry), (2) 1242 grams of a silt-clay mixture (ASP 400), (3) 77.3 grams of alpha cellulose, (4) 0.15 grams of humic

acid, and (5) 7.5 grams of dolomite (the dolomite is a source of bicarbonate buffering that occurs naturally in soils and sediments). Steps for processing the sand before use include: (1) rinsing sand with gentle mixing in well water (hardness 283 mg/L as $CaCO_3$, alkalinity 255 mg/L as $CaCO_3$, pH 7.8) until the water runs clear, (2) rinsing the sand for 5 min with deionized water, and (3) air drying the sand. Constituents are mixed for 1 h on a rolling mill and stored dry until ready for use (i.e., no conditioning required). When formulated sediments are made with a high silt-clay content, the alkalinity and hardness of the pore water may drop due to cation exchange. Gentle mixing of the formulated sediment with overlying water before use in testing reduces this change in the water quality characteristics of the pore water.

7.3 Reagents

7.3.1 Data sheets should be followed for reagents and other chemicals purchased from supply houses. The test material(s) should be at least reagent grade, unless a test using a formulated commercial product, technical-grade, or use-grade material is specifically needed. Reagent containers should be dated when received from the supplier, and the shelf life of the reagent should not be exceeded. Working solutions should be dated when prepared and the recommended shelf life should not be exceeded.

7.4 Standards

7.4.1 Appropriate standard methods for chemical and physical analyses should be used when possible. For those measurements for which standards do not exist or are not sensitive enough, methods should be obtained from other reliable sources.

Section 8
Sample Collection, Storage, Manipulation, and Characterization

8.1 Collection

8.1.1 Before the preparation or collection of sediment, a procedure should be established for the handling of sediment that might contain unknown quantities of toxic chemicals (Section 5).

8.1.2 Sediments are spatially and temporally variable (Stemmer et al., 1990a). Replicate samples should be collected to determine variance in sediment characteristics. Sediment should be collected with as little disruption as possible; however, subsampling, compositing, or homogenization of sediment samples might be necessary for some experimental designs. Sampling can cause loss of sediment integrity, change in chemical speciation, or disruption of chemical equilibrium (ASTM, 1999b). A benthic grab or core should be used rather than a dredge to minimize disruption of the sediment sample. Sediment should be collected from a depth that will represent expected exposure. For example, oligochaetes may burrow 4 to 15 cm into sediment. Samples collected for evaluations of dredged material should include sediment cores to the depth of removal. Surveys of the toxicity of surficial sediment are often based on cores of the upper 2 cm sediment depth.

8.1.3 Exposure to direct sunlight during collection should be minimized, especially if the sediment contains photolytic compounds. Sediment samples should be cooled to 4°C in the field before shipment (ASTM, 1999b). Dry ice can be used to cool samples in the field; however, sediments should never be frozen. Monitors can be used to measure temperature during shipping (e.g., TempTale Temperature Monitoring and Recording System, Sensitech, Inc., Beverly, MA).

8.1.4 For additional information on sediment collection and shipment see USEPA (1999) and ASTM (1999b).

8.2 Storage

8.2.1 Since the contaminants of concern and influencing sediment characteristics are not always known, it is desirable to hold the sediments after collection in the dark at 4°C. Traditional convention has held that toxicity tests should be started as soon as possible following collection from the field, although actual recommended storage times range from two weeks (ASTM, 1999b) to less than eight weeks (USEPA-USACE, 1998a). Discrepancies in recommended storage times reflected a lack of data concerning the effects of long-term storage on the physical, chemical, and toxicological characteristics of the sediment. However, numerous studies have recently been conducted to address issues related to sediment storage (Dillon et al., 1994; Becker and Ginn, 1995; Carr and Chapman, 1995; Moore et al., 1996; Sarda and Burton, 1995; Sijm et al., 1997; DeFoe and Ankley, 1998). The conclusions and recommendations offered by these studies vary substantially and appear to depend primarily upon the type or class of contaminant(s) present. Considered collectively, these studies suggest that the recommended guidance that sediments be tested sometime between the time of collection and 8 weeks storage is appropriate. Additional guidance is provided below.

8.2.2 Extended storage of sediments that contain high concentrations of labile chemicals (e.g., ammonia, volatile organics) may lead to a loss of these chemicals and a corresponding reduction in toxicity. Under these circumstances, the sediment should be tested as soon as possible after collection, but not later than within two weeks (Sarda and Burton, 1995). Sediments that exhibit low-level to moderate toxicity can exhibit considerable temporal variability in toxicity, although the direction of change is often unpredictable (Carr and Chapman, 1995; Moore et al., 1996; DeFoe and Ankley, 1998). For these types of sediments, the recommended storage time of <8 weeks may be most appropriate. In some situations, a minimum storage period for low-to-moderately contaminated sediments may help reduce variability. For example, DeFoe and Ankley (1998) observed high variability in survival during early testing periods (e.g., <2 weeks) in sediments with low toxicity. DeFoe and Ankley (1998) hypothesized that this variability partially reflected the presence of indigenous predators that remained alive during this relatively short storage period. Thus, if predatory species are known to exist, and the sediment does not contain labile contaminants, it may be desirable to store the sediment for a short period before testing (e.g., 2 weeks) to reduce potential for interferences from indigenous organisms. Sediments that contain comparatively stable compounds (e.g., high molecular weight compounds such as PCBs) or which exhibit a moderate-to-high level of toxicity, typically do not vary appreciably in toxicity in relation to storage duration (Moore et al., 1996; DeFoe and Ankley, 1998). For these sediments, long-term storage (e.g., >8 weeks) can be undertaken.

8.2.3 Researchers may wish to conduct additional characterizations of sediment to evaluate possible effects of storage. Concentrations of chemicals of concern could be measured periodically in pore water during the storage period and at the start of the sediment test (Kemble et al., 1994). Ingersoll et al. (1993) recommend conducting a toxicity test with pore water within two weeks from sediment collection and at the start of the sediment test. Freezing might further change sediment properties such as grain size or chemical partitioning and should be avoided (ASTM, 1999b; Schuytema et al., 1989). Sediment should be stored with no air over the sealed samples (no head space) at 4°C before the start of a test (Shuba et al.,1978). Sediment may be stored in containers constructed of suitable materials as outlined in Section 6.

8.3 Manipulation

8.3.1 Homogenization

8.3.1.1 Samples tend to settle during shipment. As a result, water above the sediment should not be discarded but should be mixed back into the sediment during homogenization. Sediment samples should not be sieved to remove indigenous organisms unless there is a good reason to believe indigenous organisms may influence the response of the test organism. However, large indigenous organisms and large debris can be removed using forceps. Reynoldson et al. (1994) observed reduced growth of amphipods, midges, and mayflies in sediments with elevated numbers of oligochaetes and recommended sieving sediments suspected to have high numbers of indigenous oligochaetes. If sediments must be sieved, it may be desirable to analyze samples before and after sieving (e.g., pore-water metals, DOC, AVS, TOC) to document the influence of sieving on sediment chemistry.

8.3.1.2 If sediment is collected from multiple field samples, the sediment can be pooled and mixed by stirring or using a rolling mill, feed mixer, or other suitable apparatus (see ASTM, 1999b). Homogenization of sediment can be accomplished using a variable-speed hand-held drill outfitted with a stainless-steel auger (diameter 7.6 cm, overall length 38 cm, auger bit length 25.4 cm; Part No. 800707, Augers Unlimited, Exton, PA; Kemble et al., 1994).

8.3.2 Sediment Spiking

8.3.2.1 Test sediment can be prepared by manipulating the properties of a control sediment. Mixing time (Stemmer et al., 1990a) and aging (Word et al., 1987; Landrum, 1989; Landrum and Faust, 1992) of spiked sediment can affect bioavailability of chemicals in sediment. Many studies with spiked sediment are often started only a few days after the chemical has been added to the sediment. This short time period may not be long enough for sediments to equilibrate with the spiked chemicals (Section 8.3.2.2.3). Consistent spiking procedures should be followed in order to make interlaboratory comparisons. See USEPA (1999) and ASTM (1999b) for additional detail regarding sediment spiking.

8.3.2.1.1 The cause of sediment toxicity and the magnitude of interactive effects of chemicals can be estimated by spiking a sediment with chemicals or complex waste mixtures (Lamberson and Swartz, 1992). Sediments spiked with a range of concentrations can be used to generate either point estimates (e.g., LC50) or a minimum concentration at which effects are observed (lowest observed effect concentration; LOEC). Results of tests may be reported in terms of a BSAF (Ankley et al., 1992b). The influence of sediment physico-chemical characteristics on chemical toxicity can also be determined with sediment-spiking studies (Adams et al., 1985).

8.3.2.2 The test material(s) should be at least reagent grade, unless a test using a formulated commercial product, technical-grade, or use-grade material is specifically needed. Before a test is started, the following should be known about the test material: (1) the identity and concentration of major ingredients and impurities, (2) water solubility in test water, (3) log Kow, BCF (from other test species), persistence, hydrolysis, and photolysis rates of the test substances, (4) estimated toxicity to the test organism and to humans, (5) if the test concentration(s) are to be measured, the precision and bias of the analytical method at the planned concentration(s) of the test material, and (6) recommended handling and disposal procedures. Addition of test material(s) to sediment may be accomplished using various methods, such as a (1) rolling mill, (2) feed mixer, or (3) hand mixing (ASTM, 1999b; USEPA, 1999). Modifications of the mixing techniques might be necessary to allow time for a test material to equilibrate with the sediment. Mixing time of spiked sediment should be limited from minutes to a few hours, and temperature should be kept low to minimize potential changes in the physico-chemical and microbial characteristics of the sediment (ASTM, 1999b). Duration of contact between the chemical and sediment can affect partitioning and bioavailability (Word et al., 1987). Care should be taken to ensure that the chemical is thoroughly and evenly distributed in the sediment. Analyses of sediment subsamples are advisable to determine the degree of mixing homogeneity (Ditsworth et al., 1990). Moreover, results from sediment-spiking studies should be compared to the response of test organisms to chemical concentrations in natural sediments (Lamberson and Swartz, 1992).

8.3.2.2.1 Organic chemicals have been added: (1) directly in a dry (crystalline) form; (2) coated on the inside walls of the container (Ditsworth et al., 1990); or (3) coated onto silica sand (e.g., 5% w/w of sediment) which is added to the sediment (D.R. Mount, USEPA, Duluth, MN, personal communication). In techniques 2 and 3, the chemical is dissolved in solvent, placed in a glass spiking container (with or without sand), then the solvent is slowly evaporated. The advantage of these three approaches is that no solvent is introduced to the sediment, only the chemical being spiked. When testing spiked sediments, procedural blanks (sediments that have been handled in the same way, including solvent addition and evaporation, but contain no added chemical) should be tested in addition to regular negative controls.

8.3.2.2.2 Metals are generally added in an aqueous solution (ASTM, 1999b; Carlson et al., 1991; Di Toro et al., 1990). Ammonia has also been successfully spiked using aqueous solutions (Besser et al., 1998). Inclusion of spiking blanks is recommended.

8.3.2.2.3 Sufficient time should be allowed after spiking for the spiked chemical to equilibrate with sediment components. For organic chemicals, it is recommended that the sediment be aged at least one month before starting a test. Two months or more may be necessary for chemicals with a high log Kow (e.g., >6; D.R. Mount, USEPA, Duluth, MN, personal communication). For metals, shorter aging times (1 to 2 weeks) may be sufficient. Periodic monitoring of chemical concentrations in pore water during sediment aging is highly recommended as a means to assess the equilibration of the spiked sediments. Monitoring of pore water during spiked sediment testing is also recommended.

8.3.2.3 Direct addition of a solvent (other than water) to the sediment should be avoided if possible. Addition of organic solvents may dramatically influence the concentration of dissolved organic carbon in pore water. If an organic solvent is to be used, the solvent should be at a concentration that does not affect the test organism. Further, both solvent control and negative control sediments must be included in the test. The solvent control must contain the highest concentration of solvent present and must be from the same batch used to make the stock solution (see ASTM, 1999e).

8.3.2.4 If the test contains both a negative control and a solvent control, the survival, growth, or reproduction of the organisms tested should be compared. If a statistically significant difference is detected between the two controls, only the solvent control may be used for meeting the acceptability of the test and as the basis for calculating results. The negative control might provide additional information on the general health of the organisms tested. If no statistically significant difference is detected, the data from both controls should be used for meeting the acceptability of the test and as the basis for calculating the results (ASTM, 1999f). If performance in the solvent control is markedly different from that in the negative control, it is possible that the data are compromised by experimental artifacts and may not accurately reflect the toxicity of the chemical in natural sediments.

8.3.3 Test Concentration(s) for Laboratory Spiked Sediments

8.3.3.1 If a test is intended to generate an LC50, a toxicant concentration series (0.5 or higher) should be selected that will provide partial mortalities at two or more concentrations of the test chemical. The LC50 of a particular compound may vary depending on physical and chemical sediment characteristics. It may be desirable to conduct a range-finding test in which the organisms are exposed to a control and three or more concentrations of the test material that differ by a factor of ten. Results from water-only tests could be used to establish concentrations to be tested in a whole-sediment test based on predicted pore-water concentrations (Di Toro et al., 1991).

8.3.3.2 Bulk-sediment chemical concentrations might be normalized to factors other than dry weight. For example, concentrations of nonpolar organic compounds might be normalized to sediment organic-carbon content, and simultaneously extracted metals might be normalized to acid volatile sulfides (Di Toro et al., 1990; Di Toro et al., 1991).

8.3.3.3 In some situations it might be necessary to simply determine whether a specific concentration of test material is toxic to the test organism, or whether adverse effects occur above or below a specific concentration. When there is interest in a particular concentration, it might only be necessary to test that concentration and not to determine an LC50.

8.4 Characterization

8.4.1 All sediments should be characterized and at least the following determined: pH and ammonia of the pore water, organic carbon content (total organic carbon, TOC), particle size distribution (percent sand, silt, clay), and percent water content (ASTM, 1999a; Plumb, 1981). See Section 8.4.4.7 for methods to isolate pore water.

8.4.2 Other analyses on sediments might include biological oxygen demand, chemical oxygen demand, cation exchange capacity, Eh, total inorganic carbon, total volatile solids, acid volatile sulfides, metals, synthetic organic compounds, oil and grease, petroleum hydrocarbons, as well as interstitial water analyses for various physicochemical parameters.

8.4.3 Macrobenthos may be evaluated by subsampling the field-collected sediment. If direct comparisons are to be made, subsamples for toxicity testing should be collected from the same sample to be used for analysis of sediment physical and chemical characterizations. Qualitative descriptions of the sediment can include color, texture, and presence of macrophytes or animals. Monitoring the odor of sediment samples should be avoided because of potential hazardous volatile chemicals.

8.4.4 Analytical Methodology

8.4.4.1 Chemical and physical data should be obtained using appropriate standard methods whenever possible. For those measurements for which standard methods do not exist or are not sensitive enough, methods should be obtained from other reliable sources.

8.4.4.2 The precision, accuracy, and bias of each analytical method used should be determined in the appropriate matrix: that is, sediment, water, tissue. Reagent blanks and analytical standards should be analyzed, and recoveries should be calculated.

8.4.4.3 Concentration of spiked test material(s) in sediment, interstitial water, and overlying water should be

measured as often as practical during a test. If possible, the concentration of the test material in overlying water, interstitial water and sediments should be measured at the start and end of a test. Measurement of test material(s) degradation products might also be desirable.

8.4.4.4 Separate chambers should be set up at the start of a test and destructively sampled during and at the end of the test to monitor sediment chemistry. Test organisms and food should be added to these extra chambers.

8.4.4.5 Measurement of test material(s) concentration in water can be accomplished by pipeting water samples from about 1 to 2 cm above the sediment surface in the test chamber. Overlying water samples should not contain any surface debris, any material from the sides of the test chamber, or any sediment.

8.4.4.6 Measurement of test material(s) concentration in sediment at the end of a test can be taken by siphoning most of the overlying water without disturbing the surface of the sediment, then removing appropriate aliquots of the sediment for chemical analysis.

8.4.4.7 Interstitial water

8.4.4.7.1 Interstitial water (pore water), defined as the water occupying the spaces between sediment or soil particles, is often isolated to provide either a matrix for toxicity testing or to provide an indication of the concentration or partitioning of contaminants within the sediment matrix. Draft USEPA equilibrium partitioning sediment guidelines (ESGs) are based on the presumption that the concentration of chemicals in the interstitial water are correlated directly to their bioavailability and, therefore, their toxicity (Di Toro et al., 1991). Of additional importance is contaminants in interstitial waters can be transported into overlying waters through diffusion, bioturbation, and resuspension processes (Van Rees et al., 1991). The usefulness of interstitial water sampling for determining chemical contamination or toxicity will depend on the study objectives and nature of the sediments at the study site.

8.4.4.7.2 Isolation of sediment interstitial water can be accomplished by a wide variety of methods, which are based on either physical separation or on diffusion/equilibrium. The common physical-isolation procedures can be categorized as: (1) centrifugation, (2) compression/squeezing, or (3) suction/vacuum. Diffusion/equilibrium procedures rely on the movement (diffusion) of pore-water constituents across semipermeable membranes into a collecting chamber until an equilibrium is established. A description of the materials and procedures used in the isolation of pore water is included in the reviews by Bufflap and Allen (1995a), ASTM (1999b), and USEPA (1999).

8.4.4.7.3 When relatively large volumes of water are required (>20 mL) for toxicity testing or chemical analyses, appropriate quantities of sediment are generally collected with grabs or corers for subsequent isolation of the interstitial water. Several isolation procedures, such as centrifugation (Ankley and Scheubauer-Berigan, 1994), squeezing (Carr and Chapman, 1995) and suction (Winger and Lasier, 1991; Winger et al., 1998), have been used successfully to obtain adequate volumes for testing purposes. Peepers (dialysis) generally do not produce sufficient volumes for most analyses; however, larger sized peepers (500-mL volume) have been used for collecting interstitial water *in situ* for chemical analyses and organism exposures (Burton, 1992; Sarda and Burton, 1995).

8.4.4.7.4 There is no one superior method for the isolation of interstitial water used for toxicity testing and associated chemical analyses. Factors to consider in the selection of an isolation procedure may include: (1) volume of pore water needed, (2) ease of isolation (materials, preparation time, and time required for isolation), and (3) artifacts in the pore water caused by the isolation procedure. Each approach has unique strengths and limitations (Bufflap and Allen, 1995a, 1995b; Winger et al., 1998), which vary with sediment characteristics, chemicals of concern, toxicity test methods, and desired test resolution (i.e., data quality objectives). For suction or compression separation, which uses a filter or a similar surface, there may be changes to the characteristics of the interstitial water compared with separation using centrifugation (Ankley et al., 1994; Horowitz et al., 1996). For most toxicity test procedures, relatively large volumes of interstitial water (e.g., liters) are frequently needed for static or renewal exposures with the associated water chemistry analyses. Although centrifugation can be used to generate large volumes of interstitial water, it is difficult to use centrifugation to isolate water from coarser sediment. If smaller volumes of interstitial water are adequate and logistics allow, it may be desirable to use peepers, which establish an equilibrium with the pore water through a permeable membrane. If logistics do not allow placement of peeper samplers, an alternative procedure could be to collect cores that can be sampled using side port suctioning or centrifugation (G.A. Burton, Wright State University, personal communication). However, if larger samples of interstitial water are needed, it would be necessary to collect multiple cores as quickly as possible using an inert environment and to centrifuge samples at ambient temperatures. See USEPA (1999) and ASTM (1999b) for additional detail regarding isolation of interstitial water.

Section 9
Quality Assurance and Quality Control

9.1 Introduction

9.1.1 Developing and maintaining a laboratory quality assurance (QA) program requires an ongoing commitment by laboratory management and also includes the following: (1) appointment of a laboratory quality assurance officer with the responsibility and authority to develop and maintain a QA program, (2) preparation of a Quality Assurance Project Plan with Data Quality Objectives, (3) preparation of written descriptions of laboratory Standard Operating Procedures (SOPs) for test organism culturing, testing, instrument calibration, sample chain-of-custody, laboratory sample tracking system, and (4) provision of adequate, qualified technical staff and suitable space and equipment to assure reliable data. Additional guidance for QA can be obtained in USEPA (1989d) and in USEPA (1994e).

9.1.2 QA practices within a testing laboratory should address all activities that affect the quality of the final data, such as (1) sediment sampling and handling, (2) the source and condition of the test organisms, (3) condition and operation of equipment, (4) test conditions, (5) instrument calibration, (6) replication, (7) use of reference toxicants, (8) record keeping, and (9) data evaluation.

9.1.3 Quality control (QC) practices, on the other hand, consist of the more focused, routine, day-to-day activities carried out within the scope of the overall QA program. For more detailed discussion of quality assurance, and general guidance on good laboratory practices related to testing see FDA (1978), USEPA (1979a), USEPA (1980a), USEPA (1980b), USEPA (1991a), USEPA (1994c), USEPA (1994d), USEPA (1995), DeWoskin (1984), and Taylor (1987).

9.2 Performance-based Criteria

9.2.1 USEPA Environmental Monitoring Management Council (EMMC) recommended the use of performance-based methods in developing standards for chemical analytical methods (Williams, 1993). Performance-based methods were defined by EMMC as a monitoring approach that permits the use of appropriate methods that meet pre-established demonstrated performance standards. Minimum required elements of performance, such as precision, reproducibility, bias, sensitivity, and detection limits should be specified, and the method should be demonstrated to meet the performance standards.

9.2.2 Participants at a September 1992 USEPA sediment toxicity workshop arrived at a consensus on several culturing and testing methods for freshwater organisms (Appendix A of USEPA, 1994a). In developing guidance for culturing test organisms to be included in this manual for sediment tests, it was generally agreed that no single method must be used to culture organisms. Success of a test relies on the health of the culture from which organisms are taken for testing. Having healthy organisms of known quality and age for testing is the key consideration relative to culture methods. Therefore, a performance-based criteria approach is the preferred method through which individual laboratories should evaluate culture health rather than using control-based criteria. Performance-based criteria were chosen to allow each laboratory to optimize culture methods while providing organisms that produce reliable and comparable test results. See Tables 11.3, 12.3, 13.4, 14.3 and 15.3 for a listing of performance criteria for culturing and testing.

9.3 Facilities, Equipment, and Test Chambers

9.3.1 Separate areas for test organism culturing and testing must be provided to avoid loss of cultures due to cross-contamination. Ventilation systems should be designed and operated to prevent recirculation or leakage of air from chemical analysis laboratories or sample storage and preparation areas into test organism culturing or sediment testing areas, and from sediment testing laboratories and sample preparation areas into culture rooms.

9.3.2 Equipment for temperature control should be adequate to maintain recommended test-water temperatures. Recommended materials should be used in the fabricating of the test equipment that comes in contact with the sediment or overlying water.

9.3.3 Before a sediment test is conducted in a new facility, a “noncontaminant” test should be conducted in which all test chambers contain a control sediment and overlying water. This information is used to demonstrate that the facility, control sediment, water, and handling procedures provide acceptable responses of test organisms (See Section 9.14).

9.4 Test Organisms

9.4.1 The organisms should appear healthy, behave normally, feed well, and have low mortality in cultures, during holding (e.g., <20% for 48 h before the start of a test), and in test controls. The species of test organisms should be positively identified to species.

9.5 Water

9.5.1 The quality of water used for organism culturing and testing is extremely important. Overlying water used in testing and water used in culturing organisms should be uniform in quality. Acceptable water should allow satisfactory survival, growth, or reproduction of the test organisms. Test organisms should not show signs of disease or apparent stress (e.g., discoloration, unusual behavior). See Section 7 for additional details.

9.6 Sample Collection and Storage

9.6.1 Sample holding times and temperatures should conform to conditions described in Section 8.

9.7 Test Conditions

9.7.1 It is desirable to measure temperature continuously in at least one chamber during each test. Temperatures should be maintained within the limits specified for each test. Dissolved oxygen, alkalinity, water hardness, conductivity, ammonia, and pH should be checked as prescribed in Sections 11.3, 12.3, 13.3, 14.3 and 15.3.

9.8 Quality of Test Organisms

9.8.1 It may be desirable for laboratories to periodically perform 96-h water-only reference-toxicity tests to assess the sensitivity of culture organisms (Section 9.16). Data from these reference-toxicity tests could be used to assess genetic strain or life-stage sensitivity to select chemicals. The requirement in the first edition of this manual for laboratories to conduct monthly reference-toxicity tests (USEPA, 1994a) has not been included as a requirement in this second edition for testing sediments because of the inability of reference-toxicity tests to identify stressed populations of test organisms (McNulty et al., 1999). Physiological measurements such as lipid content might also provide useful information regarding the health of the cultures.

9.8.2 It is desirable to determine the sensitivity of test organisms obtained from an outside source. The supplier should provide data with the shipment describing the history of the sensitivity of organisms from the same source culture. The supplier should also certify the species identification of the test organisms and provide the taxonomic references or name(s) of the taxonomic expert(s) consulted.

9.8.3 All organisms in a test must be from the same source (Section 10.2.2). Organisms may be obtained from laboratory cultures or from commercial or government sources (Table 10.1). The test organisms used should be identified using an appropriate taxonomic key, and verification should be documented (Pennak, 1989; Merritt and Cummins, 1996). Obtaining organisms from wild populations should be avoided unless organisms are cultured through several generations in the laboratory. In addition, the ability of the wild population of sexually reproducing organisms to cross breed with the existing laboratory population should be determined (Duan et al.,1997). Sensitivity of the wild population to select chemicals (e.g., Table 1.4) should also be documented.

9.9 Quality of Food

9.9.1 Problems with the nutritional suitability of the food will be reflected in the survival, growth, or reproduction of the test organisms in cultures or in sediment tests.

9.9.2 Food used to culture organisms used in bioaccumulation tests must be analyzed for compounds to be measured in the bioaccumulation tests.

9.10 Test Acceptability

9.10.1 Tables 11.3, 12.3, 13.4, 14.3 and 15.3 outline requirements for acceptability of the tests. An individual test may be conditionally acceptable if temperature, dissolved oxygen, and other specified conditions fall outside specifications, depending on the degree of the departure and the objectives of the tests (see test condition summaries in Tables 11.1, 12.1, 13.1, 14.1, and 15.1). The acceptability of a test will depend on the experience and professional judgment of the laboratory analyst and the reviewing staff of the regulatory authority. Any deviation from test specifications should be noted when reporting data from a test.

9.11 Analytical Methods

9.11.1 All routine chemical and physical analyses for culture and testing water, food, and sediment should include established quality assurance practices outlined in USEPA methods manuals (USEPA, 1979a; USEPA, 1979b; USEPA, 1991a; USEPA, 1994b).

9.11.2 Reagent containers should be dated when received from the supplier, and the shelf life of the reagent should not be exceeded. Working solutions should be dated when prepared and the recommended shelf life should not be exceeded.

9.12 Calibration and Standardization

9.12.1 Instruments used for routine measurements of chemical and physical characteristics such as pH, dissolved oxygen, temperature, and conductivity should be calibrated before use each day according to the instrument manufacturer's procedures as indicated in the general section on quality assurance (see USEPA Methods 150.1, 360.1, 170.1, and 120.1; USEPA, 1979b). Calibration data should be recorded in a permanent log.

9.12.2 A known-quality water should be included in the analyses of each batch of water samples (e.g., water hardness, alkalinity, conductivity). It is desirable to include certified standards in the analysis of water samples.

9.13 Replication and Test Sensitivity

9.13.1 The sensitivity of sediment tests will depend in part on the number of replicates/treatment, the significance level selected, and the type of statistical analysis. If the variability remains constant, the sensitivity of a test will increase as the number of replicates is increased. The minimum recommended number of replicates varies with the objectives of the test and the statistical method used for analysis of the data (Section 16).

9.14 Demonstrating Acceptable Performance

9.14.1 Intralaboratory precision, expressed as a coefficient of variation (CV) of the range in response for each type of test to be used in a laboratory, can be determined by performing five or more tests with different batches of test organisms using the same reference toxicant at the same concentrations with the same test conditions (e.g., the same test duration, type of water, age of test organisms, feeding) and the same data analysis methods. This should be done to gain experience for the toxicity tests and to serve as a point of reference for future testing. A reference-toxicity concentration series (0.5 or higher) should be selected that will provide partial mortalities at two or more concentrations of the test chemical (Section 8.3.3). Information from previous tests can be used to improve the design of subsequent tests to optimize the dilution series selected for future testing.

9.14.2 Before conducting tests with potentially contaminated sediment, it is strongly recommended that the laboratory conduct the tests with control sediment(s) alone. Results of these preliminary studies should be used to determine if use of the control sediment and other test conditions (i.e., water quality) result in acceptable performance in the tests as outlined in Tables 11.1, 12.1, 13.1, 14.1, and 15.1.

9.14.3 Laboratories should demonstrate that their personnel are able to recover an average of at least 90% of the organisms from whole sediment. For example, test organisms could be added to control sediment or test sediments and recovery could be determined after 1 h (Tomasovic et al., 1994).

9.15 Documenting Ongoing Laboratory Performance

9.15.1 Outliers, which are data falling outside the control limits, and trends of increasing or decreasing sensitivity are readily identified. If the reference-toxicity results from a given test fall outside the "expected" range (e.g., +2 SD), the sensitivity of the organisms and the credibility of the test results may be suspect. In this case, the test procedure should be examined for defects and should be repeated with a different batch of test organisms (Section 16).

9.15.2 A sediment test may be acceptable if specified conditions of a reference-toxicity test fall outside the expected ranges (Section 9.10.2). Specifically, a sediment test should not be judged unacceptable if the LC50 for a given reference-toxicity test falls outside the expected range or if mortality in the control of the reference-toxicity test exceeds 10%. All the performance criteria outlined in Tables 11.3, 12.3, 13.4, 14.3, and 15.3 must be considered when determining the acceptability of a sediment test. The acceptability of the sediment test would depend on the experience and judgment of the investigator and the regulatory authority.

9.15.3 Performance should improve with experience, and the control limits should gradually narrow, as the statistics stabilize. However, control limits of +2 SD, by definition, will be exceeded 5% of the time, regardless of how well a laboratory performs. For this reason, good laboratories that develop very narrow control limits may be penalized if a test result that falls just outside the control limits is rejected *de facto*. The width of the control limits should be considered in decisions regarding rejection of data (Section 17).

9.16 Reference Toxicants

9.16.1 Historically, reference-toxicity testing has been thought to provide three types of information relevant to the interpretation of toxicity test data: (1) an indication of the relative "health" of the organisms used in the test; (2) a demonstration that the laboratory can perform the test procedure in a reproducible manner; and (3) information to indicate whether the sensitivity of the particular strain or population in use at a laboratory is comparable to those in use in other facilities. With regard to the first type of information, recent work by McNulty et al. (1999) suggests that reference-toxicity tests may not be effective in identifying stressed populations of test organisms. In addition, reference-toxicity tests recommended for use with sediment toxicity tests are short-term, water column tests, owing in part to the lack of a standard sediment for reference-toxicity testing. Because the test procedures for reference-toxicity tests are not the same as for the sediment toxicity tests of interest, the applicability of reference-toxicity tests to demonstrate ability to reproducibly perform the sediment test procedures is greatly reduced. Particularly for the long-term sediment toxicity tests with *H. azteca* and *C. tentans*, performance of control organisms over time may be a better indicator of success in handling and testing these organisms (Sections 14 and 15).

9.16.2 Although the requirement for monthly testing has been removed in this second edition of the manual, periodic reference-toxicity testing should still be conducted as an indication of overall comparability of results among laboratories (at a minimum, six tests over a 3-year period should be conducted to evaluate potential differences in life stage or genetic strain of test organisms). In

particular, reference-toxicity tests should be performed more frequently when organisms are obtained from outside sources, when there are changes in culture practices, or when brood stock from an outside source is incorporated into a laboratory culture.

9.16.3 Reference toxicants such as sodium chloride (NaCl), potassium chloride (KCl), cadmium chloride ($CdCl_2$), and copper sulfate ($CuSO_4$) are suitable for use. No one reference toxicant can be used to measure the sensitivity of test organisms with respect to another toxicant with a different mode of action (Lee, 1980). However, it may be unrealistic to test more than one or two reference toxicants routinely. KCl has been used successfully in round-robin water-only exposures with *H. azteca* and *C. tentans* (Section 17).

9.16.4 Test conditions for conducting reference-toxicity tests with *H. azteca, C. tentans,* and *L. variegatus* are outlined in Tables 9.1 and 9.2. Reference-toxicity tests can be conducted using one organism/chamber or multiple organisms in each chamber. Some laboratories have observed low control survival when more than one midge/chamber is tested in water-only exposures.

9.17 Record Keeping

9.17.1 Section 16.1 outlines recommendations for record keeping (i.e., data files, chain-of-custody).

Table 9.1 Recommended Test Conditions for Conducting Reference-toxicity Tests with One Organism/Chamber

Parameter	Conditions
1. Test type:	Water-only test
2. Dilution series:	Control and at least 5 test concentrations (0.5 dilution factor)
3. Toxicant:	NaCl, KCl, Cd, or Cu
4. Temperature:	23 ± 1°C
5. Light quality:	Wide-spectrum fluorescent lights
6. Illuminance:	About 100 to 1000 lux
7. Photoperiod:	16L:8D
8. Renewal of water:	None
9. Age of organisms:	*H. azteca:* 7- to 14-d old (1- to 2-d range in age) *C. tentans:* second- to third-instar larvae (about 10-d-old larvae)[1] *L. variegatus:* adults
10. Test chamber:	30-mL plastic cups (covered with glass or plastic)
11. Volume of water:	20 mL
12. Number of organisms/chamber:	1
13. Number of replicate chambers/treatment:	10 minimum
14. Feeding:	*H. azteca:* 0.1 mL YCT (1800 mg/L stock) on Day 0 and 2 *C. tentans:* 0.25 mL Tetrafin® (4 g/L stock) on Day 0 and 2 *L. variegatus:* not fed
15. Substrate:	*H. azteca:* Nitex® screen (110 mesh) *C. tentans:* sand (monolayer) *L. variegatus:* no substrate
16. Aeration:	None
17. Dilution water:	Culture water, well water, surface water, site water, or reconstituted water
18. Test chamber cleaning:	None
19. Water quality:	Hardness, alkalinity, conductivity, dissolved oxygen, and pH at the beginning and end of a test. Temperature daily.
20. Test duration:	96 h
21. Endpoint:	Survival (LC50)
22. Test acceptability:	90% control survival

[1] Age requirement: All animals must be third or second instar with at least 50% of the organisms at third instar.

Table 9.2 Recommended Test Conditions for Conducting Reference-toxicity Tests with More Than One Organism/Chamber

	Parameter	Conditions
1.	Test type:	Water-only test
2.	Dilution series:	Control and at least 5 test concentrations (0.5 dilution factor)
3.	Toxicant:	NaCl, KCl, Cd, or Cu
4.	Temperature:	23 ± 1°C
5.	Light quality:	Wide-spectrum fluorescent lights
6.	Illuminance:	About 100 to 1000 lux
7.	Photoperiod:	16L:8D
8.	Renewal of water:	None
9.	Age of organisms:	*H. azteca:* 7- to 14-d old (1- to 2-d range in age) *C. tentans:* second to third instar (about 10-d-old larvae)[1] *L. variegatus:* adults
10.	Test chamber:	250-mL glass beaker (covered with glass or plastic)
11.	Volume of water:	100 mL (minimum)
12.	Number of organisms/chamber:	10 minimum
13.	Number of replicate chambers/treatment:	3 minimum
14.	Feeding:	*H. azteca:* 0.5 mL YCT (1800 mg/L stock) on Day 0 and 2 *C. tentans:* 1.25 mL Tetrafin® (4 g/L stock) on Day 0 and 2 *L. variegatus:* not fed
15.	Substrate:	*H. azteca:* Nitex® screen (110 mesh) *C. tentans:* sand (monolayer) *L. variegatus:* no substrate
16.	Aeration:	None
17.	Dilution water:	Culture water, well water, surface water, site water or reconstituted water
18.	Test chamber cleaning:	None
19.	Water quality:	Hardness, alkalinity, conductivity, dissolved oxygen, and pH at the beginning and end of a test. Temperature daily.
20.	Test duration:	96 h
21.	Endpoint:	Survival (LC50)
22.	Test acceptability:	90% control survival

[1] Age requirement: All animals must be third or second instar with at least 50% of the organisms at third instar.

Section 10
Collecting, Culturing, and Maintaining Test Organisms

10.1 Life Histories

10.1.1 *Hyalella azteca*

10.1.1.1 *Hyalella azteca* inhabit permanent lakes, ponds, and streams throughout North and South America (de March, 1981; Pennak, 1989). Occurrence of *H. azteca* is most common in warm (20°C to 30°C for much of the summer) mesotrophic or eutrophic lakes that support aquatic plants. These amphipods are also found in ponds, sloughs, marshes, rivers, ditches, streams, and springs, but in lower numbers. *Hyalella azteca* have achieved densities of >10,000/m^2 in preferred habitats (de March, 1981).

10.1.1.2 *Hyalella azteca* are epibenthic detritivores that burrow into the sediment surface. Hargrave (1970a) reported that *H. azteca* selectively ingest bacteria and algae. The behavior and feeding habits of *H. azteca* make them excellent test organisms for sediment assessments.

10.1.1.3 Reproduction by *H. azteca* is sexual. The adult males are larger than females and have larger second gnathopods (de March, 1981). Males pair with females by grasping the females (amplexus) with their gnathopods while on the backs of the females. After feeding together for 1 to 7 d the female is ready to molt and the two organisms separate for a short time while the female sheds her old exoskeleton. Once the exoskeleton is shed, the two organisms reunite and copulation occurs. The male places sperm near the marsupium of the female and her pleopods sweep the sperm into the marsupium. The organisms again separate and the female releases eggs from her oviducts into the marsupium where they are fertilized. *Hyalella azteca* average about 18 eggs/brood (Pennak, 1989) with larger organisms having more eggs (Cooper, 1965).

10.1.1.4 The developing embryos and newly hatched young are kept in the marsupium until the next molt. At 24°C to 28°C, hatching ranges from 5 to 10 d after fertilization (Embody, 1911; Bovee, 1950; Cooper, 1965). The time between molts for females is 7 to 8 d at 26°C to 28°C (Bovee, 1950). Therefore, about the time embryos hatch, the female molts and releases the young. *Hyalella azteca* average 15 broods in 152 d (Pennak, 1989). Pairing of the sexes is simultaneous with embryo incubation of the previous brood in the marsupium. *Hyalella azteca* have a minimum of nine instars (Geisler, 1944). There are 5 to 8 pre-reproductive instars (Cooper, 1965) and an indefinite number of post-reproductive instars. The first five instars form the juvenile stage of development, instar stages 6 and 7 form the adolescent stage when sexes can be differentiated, instar stage 8 is the nuptial stage, and all later instars are the adult stages of development (Pennak, 1989).

10.1.1.5 *Hyalella azteca* have been successfully cultured at illuminance of about 100 to 1000 lux (Ingersoll and Nelson, 1990; Ankley et al., 1991a; Ankley et al., 1991b). *Hyalella azteca* avoid bright light, preferring to hide under litter and feed during the day.

10.1.1.6 Temperatures tolerated by *H. azteca* range from 0 to 33°C (Embody, 1911; Bovee, 1949; Sprague, 1963). At temperatures less than 10°C the organisms rest and are immobile (de March, 1977; de March, 1978). At temperatures of 10°C to 18°C, reproduction can occur. Juveniles grow more slowly at colder temperatures and become larger adults. Smaller adults with higher reproduction are typical when organisms are grown at 18°C to 28°C. The highest rates of reproduction occur at 26°C to 28°C (de March, 1978) while lethality occurs at 33°C to 37°C (Bovee, 1949; Sprague, 1963).

10.1.1.7 *Hyalella azteca* are found in waters of widely varying types. *Hyalella azteca* can inhabit saline waters up to 29 ‰; however, their distribution in these saline waters has been correlated to water hardness (Ingersoll et al., 1992). *Hyalella azteca* inhabit water with high Mg concentrations at conductivities up to 22,000 µS/cm, but only up to 12,000 µS/cm in Na-dominated waters (Ingersoll et al., 1992). De March (1981) reported *H. azteca* were not collected from locations where calcium was less than 7 mg/L. *Hyalella azteca* have been cultured in reconstituted salt water with a salinity up to 15‰ (Ingersoll et al., 1992; Winger and Lasier, 1993). In laboratory studies, Sprague (1963) reported a 24-h LC50 for dissolved oxygen at 20°C of 0.7 mg/L. Pennak and Rosine (1976) reported similar findings. Nebeker et al. (1992) reported 48-h and 30-d LC50s for *H. azteca* of less than 0.3 mg/L dissolved oxygen. Weight and reproduction of *H. azteca* were reduced after 30-d exposure to 1.2 mg/L dissolved oxygen.

10.1.1.8 *Hyalella azteca* tolerate a wide range of substrates. Ingersoll et al. (1996) reported that *H. azteca* tolerated sediments ranging from more than 90% silt- and clay-sized particles to 100% sand-sized particles without detrimental effects on either survival or growth. *Hyalella azteca* tolerated a wide range in grain size and organic matter in 10- to 42-d tests with formulated sediment (Suedel and Rodgers, 1994; Ingersoll et al., 1998). Ankley et al. (1994a) evaluated the effects of natural sediment physico-chemical characteristics on the results of 10-d laboratory toxicity tests with *H. azteca, C. tentans,* and *L. variegatus*. Tests were conducted with and without the addition of exogenous food. Survival of organisms was decreased in tests without added food. Physico-chemical sediment characteristics including grain size and TOC were not significantly correlated to the response of *H. azteca* in either fed or unfed tests. See Sections 4.2.3 and 14.4 for additional detail regarding studies of the influence of grain size in long-term sediment toxicity tests with *H. azteca*.

10.1.2 *Chironomus tentans*

10.1.2.1 *Chironomus tentans* have a holarctic distribution (Townsend et al., 1981) and are commonly found in eutrophic ponds and lakes (Flannagan, 1971; Driver, 1977). Midge larvae are important in the diet of fish and waterfowl (Sadler, 1935; Siegfried, 1973; Driver et al., 1974; McLarney et al., 1974). Larvae of *C. tentans* usually penetrate a few cm into sediment. In both lotic and lentic habitats with soft bottoms, about 95% of the chironomid larvae occur in the upper 10 cm of substrates, and very few larvae are found below 40 cm (Townsend et al., 1981). Larvae were found under the following conditions in British Columbia lakes by Topping (1971): particle size <0.15 mm to 2.0 mm, temperature 0 to 23.3°C, dissolved oxygen 0.22 to 8.23 mg/L, pH 8.0 to 9.2, conductivity 481 to 4,136 µmhos/cm, and sediment organic carbon 1.9 to 15.5%. Larvae were absent from lakes if hydrogen sulfide concentration in overlying water exceeded 0.3 mg/L. Abundance of larvae was positively correlated with conductivity, pH, amount of food, percentages of particles in the 0.59 to 1.98 mm size range, and concentrations of Na, K, Mg, Cl, SO_4, and dissolved oxygen. Others (e.g., Curry, 1962; Oliver, 1971) have reported a temperature range of 0 to 35°C and a pH range of 7 to 10.

10.1.2.2 *Chironomus tentans* are aquatic during the larval and pupal stages. The life cycle of *C. tentans* can be divided into four distinct stages: (1) an egg stage, (2) a larval stage, consisting of four instars, (3) a pupal stage, and (4) an adult stage. Mating behavior has been described by Sadler (1935) and others (ASTM, 1999a). Males are easily distinguished from females because males have large, plumose antennae and a much thinner abdomen with visible genitalia. The male has paired genital claspers on the posterior tip of the abdomen (Townsend et al., 1981). The adult female weighs about twice as much as the male, with about 30% of the female weight contributed by the eggs. After mating, adult females oviposit a single transparent, gelatinous egg mass directly into the water. At the USEPA Office of Research and Development Laboratory (Duluth, MN), the females oviposit eggs within 24 h after emergence. Egg cases contain a variable number of eggs from about 500 to 2000 eggs/eggcase (J. Jenson, ILS, Duluth, MN, personal communication) and will hatch in 2 to 4 d at 23°C. Under optimal conditions larvae will pupate and emerge as adults after about 21 d at 23°C. Larvae begin to construct tubes (or cases) on the second or third day after hatching. The cases lengthen and enlarge as the larvae grow with the addition of small particles bound together with threads from the mouths of larvae (Sadler, 1935). The larvae draw food particles inside the tubes and also feed in the immediate vicinity of either end of the open-ended tubes with their caudal extremities anchored within the tube. The four larval stages are followed by a black-colored pupal stage (lasting about 3 d) and emergence to a terrestrial adult (imago) stage. The adult stage lasts for 3 to 5 d, during which time the adults mate during flight and the females oviposit their egg cases (2 to 3 d post-emergence; Sadler, 1935).

10.1.2.3 Grain size tolerance of *C. tentans* in sediment testing is described in Section 12.4.3 for 10-d exposures and in Section 15.4.3 for long-term exposures.

10.1.3 *Lumbriculus variegatus*

10.1.3.1 *Lumbriculus variegatus* inhabit a variety of sediment types throughout the United States and Europe (Chekanovskaya, 1962; Cook, 1969; Spencer, 1980; Brinkhurst, 1986). *Lumbriculus variegatus* typically tunnel in the upper aerobic zone of sediments of reservoirs, rivers, lakes, ponds, and marshes. When not tunneling, they bury their anterior portion in sediment and undulate their posterior portion in overlying water for respiratory exchange.

10.1.3.2 Adults of *L. variegatus* can reach a length of 40 to 90 mm, diameter of 1.0 to 1.5 mm, and wet weight of 5 to 12 mg (Call et al., 1991; Phipps et al., 1993). Lipid content is about 1.0% (wet weight, Ankley et al., 1992b; Brunson et al., 1993; Brunson et al., 1998). *Lumbriculus variegatus* most commonly reproduce asexually, although sexual reproduction has been reported (Chekanovskaya, 1962). Newly hatched worms have not been observed in cultures (Call et al., 1991; Phipps et al., 1993). Cultures consist of adults of various sizes. Populations of laboratory cultures double (number of organisms) every 10 to 14 d at 20°C (Phipps et al., 1993).

10.1.3.3 *Lumbriculus variegatus* tolerate a wide range of substrates. Ankley et al. (1994a) evaluated the effects of natural sediment physico-chemical characteristics on the results of 10-d laboratory toxicity tests with *H. azteca, C. tentans,* and *L. variegatus.* Tests were conducted with and without the addition of exogenous food. Survival and reproduction of organisms was decreased in tests without added food. Physico-chemical sediment characteristics including grain size and TOC were not significantly correlated to reproduction or growth of *L. variegatus* in either fed or unfed tests.

10.2 General Culturing Procedures

10.2.1 Acceptability of a culturing procedure is based in part on performance of organisms in culture and in the sediment test (Section 1.4 and 9.2). No single technique for culturing test organisms is required. What may work well for one laboratory may not work as well for another laboratory. While a variety of culturing procedures are outlined in Section 10.3 for *H. azteca,* in Section 10.4 for *C. tentans,* and in Section 10.5 for *L. variegatus,* organisms must meet the test acceptability requirements listed in Tables 11.3, 12.3, 13.4, 14.3, and 15.3.

10.2.2 All organisms in a test must be from the same source. Organisms may be obtained from laboratory cultures or from commercial or government sources (Table 10.1). The test organism used should be identified using an appropriate taxonomic key, and verification should be documented. Obtaining organisms from wild populations should be avoided unless organisms are cultured through several generations in the laboratory. In addition, the ability of the wild population of sexually reproducing organisms to crossbreed with the existing laboratory population should be determined (Duan et al. ,1997). Sensitivity of the wild population to select chemicals (e.g., Table 1.4) should also be documented.

Table 10.1 Sources of Starter Cultures of Test Organisms

Source	Species
U.S. Environmental Protection Agency Mid-Continent Ecological Division 6201 Congdon Boulevard Duluth, MN 55804 Teresa Norberg-King (218/529-5163, fax -5003) email: norberg-king.teresa@epa.gov	*H. azteca* *C. tentans* *L. variegatus*
U.S. Environmental Protection Agency Environmental Monitoring System Laboratory 26 W. Martin Luther Dr. Cincinnati, OH 45244 Jim Lazorchak (513/569-7076, fax -7609) email: lazorchak.jim@epa.gov	*H. azteca* *L. variegatus*
Columbia Environmental Research Center U.S. Geological Survey 4200 New Haven Road Columbia, MO 65201 Eugene Greer (573/876-1820, fax -1896) email: eugene_greer@usgs.gov	*H. azteca* *C. tentans* *L. variegatus*
Great Lakes Environmental Research Laboratory, NOAA 2205 Commonwealth Boulevard Ann Arbor, MI 48105-1593 Peter Landrum (313/741-2276, fax -2055) email: landrum@glerl.noaa.gov	*L. variegatus*
Wright State University Institute for Environmental Quality Dayton, OH 45435 Allen Burton (937/775-2201, fax -4997) email: aburton@wright.edu	*H. azteca* *C. tentans* *L. variegatus*
Michigan State University Department of Fisheries and Wildlife No. 13 Natural Resources Building East Lansing, MI 48824-1222 John Giesy (517/353-2000, fax 517/432-1984) email: jgiesy@aol.com	*H. azteca* *C. tentans* *L. variegatus*

10.2.3 Test organisms obtained from commercial sources should be shipped in well-oxygenated water in insulated containers to maintain temperature during shipment. Temperature and dissolved oxygen of the water in the shipping containers should be measured on arrival to determine if the organisms might have been subjected to low dissolved oxygen or temperature fluctuations. The temperature of the shipped water should be gradually adjusted to the desired culture temperature at a rate not exceeding 2°C per 24 h. Additional reference-toxicity testing is suggested if organisms are not cultured at the testing laboratory (Section 9.16).

10.2.4 A group of organisms should not be used for a test if they appear to be unhealthy, discolored, or otherwise stressed (e.g., >20% mortality for 48 h before the start of a test). If the organisms fail to meet these criteria, the entire batch should be discarded and a new batch should be obtained. All organisms should be as uniform as possible in age and life stage. Test organisms should be handled as little as possible. When handling is necessary, it should be done as gently, carefully, and as quickly as possible.

10.2.5 *H. azteca, C. tentans,* and *L. variegatus* can be cultured in a variety of waters. Water of a quality sufficient to culture fathead minnows (*Pimephales promelas*) or cladocerans will generally be adequate.

10.2.5.1 Variable success has been reported using reconstituted waters to culture or test *H. azteca* in long-term exposures (i.e., >10 d; See Section 7.1.3 for details).

10.2.5.2 Organisms can be cultured using either static or renewal procedures. Renewal of water is recommended to limit loss of the culture organisms from a drop in dissolved oxygen or a buildup of waste products. In renewal systems, there should be at least one volume addition/d of culture water to each chamber. In static systems, the overlying water volume should be changed at least weekly by siphoning down to a level just above the substrate and slowly adding fresh water. Extra care should be taken to ensure that proper water quality is maintained in static systems. For example, aeration is needed in static systems to maintain dissolved oxygen at >2.5 mg/L.

10.2.5.3 A recirculating system using an under-gravel filter has been used to culture amphipods and midges (P.V. Winger, USGS, Athens, GA, personal communication). The approach for using a recirculating system to culture organisms has been described by New et al. (1974), Crandall et al. (1981), and Rottmann and Campton (1989). Under-gravel filters can be purchased from aquarium suppliers and consist of an elevated plate with holes that fit on the bottom of an aquarium. The plate has a standpipe to which a pump can be attached. Gravel or

an artificial substrate (e.g., plastic balls or multi-plate substrates) is placed on the plate. The substrates provide surface area for microorganisms that use nitrogenous compounds. A simple example of a recirculating system is two aquaria positioned one above the other with a total volume of 120 L. The bottom aquarium contains the under-gravel filter system, gravel, or artificial substrate, and a submersible pump. The top aquarium is used for culture of animals and has a hole in the bottom with a standpipe for returning overflow water to the bottom aquarium. Water lost to evaporation is replaced weekly, and water is replaced at one- to two-month intervals. Cultures fed foods such as Tetramin® or Tetrafin® should include limestone gravel to help avoid depression in pH. Recirculating systems require less maintenance than static systems.

10.2.6 Cultures should be maintained at 23°C with a 16L:8D photoperiod at an illuminance of about 100 to 1000 lux (USEPA, 1994a; ASTM, 1999a). Cultures should be observed daily. Water temperature should be measured daily or continuously, and dissolved oxygen should be measured weekly. It may be desirable for laboratories to periodically perform 96-h water-only reference-toxicity tests to assess the sensitivity of culture organisms (Section 9.16.2). Data from these reference-toxicity tests could be used to assess genetic strain or life-stage sensitivity to select chemicals. The previous requirement for laboratories to conduct monthly reference-toxicity tests (USEPA 1994a) has not been included as a requirement in this second edition for testing sediments due to the inability of reference-toxicity tests to identify stressed populations of test organisms (Section 9.16; McNulty et al., 1999). Culture water hardness, alkalinity, ammonia, and pH should be measured at least quarterly. If amphipods are cultured using static conditions, it is desirable to measure water quality more frequently. If reconstituted water is used to culture organisms, water quality should be measured on each batch of reconstituted water. Culture procedures should be evaluated and adjusted as appropriate to restore or maintain the health of the culture.

10.3 Culturing Procedures for *Hyalella azteca*

10.3.1 The culturing procedures described below are based on methods described in USEPA (1991a), Ankley et al. (1994a), Call et al. (1994), Tomasovic et al. (1994), Greer (1993), Ingersoll and Nelson (1990), Ingersoll et al. (1998), ASTM (1999a) and USEPA (1994a). The culturing procedure must produce 7- to 14-d-old amphipods to start a 10-d sediment test (Table 11.3). The 10-d test with *H. azteca* should start with a narrow range in size or age of *H. azteca* (1- to 2-d range in age) to reduce potential variability in growth at the end of the 10-d test. This narrower range would be easiest to obtain using known-age organisms (i.e., Section 10.3.2, 10.3.4) instead of sieving the cultures (Section 10.3.5) to obtain similar-sized amphipods (i.e., amphipods within a range of 1- to 2-d old will be more uniform in size than organisms within the range of 7 d). The culturing procedure must produce 7- to 8-d-old amphipods to start a long-term test with *H. azteca* (Table 14.3).

10.3.2 The following procedure described by Call et al. (1994) and USEPA (1991a) can be used to obtain known-age amphipods to start a test. Mature amphipods (50 organisms >30-d old at 23°C) are held in 2-L glass beakers containing 1 L of aerated culture water and cotton gauze as a substrate. Amphipods are fed 10 mL of a yeast-Cerophyl®-trout chow (YCT) mixture (Appendix B) and 10 mL of the green algae *Selenastrum capricornutum* (about 3.5 x 10^7 cells/mL). Five mL of each food is added to each culture daily, except for renewal days, when 10 mL of each food is added.

10.3.2.1 Water in the culture chambers is changed weekly. Survival of adults and juveniles and production of young amphipods should be measured at this time. The contents of the culture chambers are poured into a translucent white plastic or white enamel pan. After the adults are removed, the remaining amphipods will range in age from <1- to 7-d old. Young amphipods are transferred with a pipet into a 1-L beaker containing culture water and are held for one week before starting a toxicity test. Organisms are fed 10 mL of YCT and 10 mL of green algae on start-up day, and 5 mL of each food each following day (Appendix B). Survival of young amphipods should be >80% during this one-week holding period. Records should be kept on the number of surviving adults, number of breeding pairs, and young production and survival. This information can be used to develop control charts that are useful in determining whether cultures are maintaining a vigorous reproductive rate indicative of culture health. Some of the adult amphipods can be expected to die in the culture chambers, but mortality greater than about 50% should be cause for concern. Reproductive rates in culture chambers containing 60 adults can be as high as 500 young per week. A decrease in reproductive rate may be caused by a change in water quality, temperature, food quality, or brood stock health. Adult females will continue to reproduce for several months.

10.3.3 A second procedure for obtaining known-age amphipods is described by Borgmann et al. (1989). Known-age amphipods are cultured in 2.5-L chambers containing about 1 L of culture water and between 5 and 25 adult *H. azteca.* Each chamber contains pieces of cotton gauze presoaked in culture water. Once a week the test organisms are isolated from the gauze and collected using a sieve. Amphipods are then rinsed into petri dishes where the young and adults are sorted. The adults are returned to the culture chambers containing fresh water and food.

10.3.4 A third procedure for obtaining known-age amphipods is described by Greer (1993), Tomasovic et al. (1994), and Ingersoll et al. (1998). Mass cultures of mixed-age amphipods are maintained in 80-L glass aquaria containing about 50 L of water (Ingersoll and Nelson, 1990). A flaked food (e.g., Tetrafin®) is added to each culture chamber receiving daily water renewals to provide about 20 g dry solids/50 L of water twice weekly in an 80-L culture chamber. Additional flaked food is added when

most of the flaked food has been consumed. Laboratories using static systems should develop lower feeding rates specific to their systems. Each culture chamber has a substrate of maple leaves and artificial substrates (six 20-cm diameter sections per 80-L aquaria of nylon "coiled-web material"; 3-M, St. Paul, MN). Before use, leaves are soaked in 30‰ salt water for about 30 d to reduce the occurrence of planaria, snails, or other organisms in the substrate. The leaves are then flushed with water to remove the salt water and residuals of naturally occurring tannic acid before placement in the cultures.

10.3.4.1 To obtain known-age amphipods, a U.S. Standard Sieve #25 (710-µm mesh) is placed underwater in a chamber containing mixed-age amphipods. A #25 sieve will retain mature amphipods, and immature amphipods will pass through the mesh. Two or three pieces of artificial substrate (3-M coiled-web material) or a mass of leaves with the associated mixed-age amphipods are quickly placed into the sieve. The sieve is brought to the top of the water in the culture chamber keeping all but about 1 cm of the sieve under water. The artificial substrates or leaves are then shaken under water several times to dislodge the attached amphipods. The artificial substrates or leaves are taken out of the sieve and placed back in the culture chamber. The sieve is agitated in the water to rinse the smaller amphipods back into the culture chamber. The larger amphipods remaining in the sieve are transferred with a pipet into a dish and then placed into a shallow glass pan (e.g., pie pan) where immature amphipods are removed. The remaining mature amphipods are transferred using a pipet into a second #25 sieve which is held in a glass pan containing culture water.

10.3.4.2 The mature amphipods are left in the sieve in the pan overnight to collect any newborn amphipods that are released. After 24 h, the sieve is moved up and down several times to rinse the newborn amphipods (<24-h old) into the surrounding water in the pan. The sieve is removed from the pan, and the mature amphipods are placed back into their culture chamber or placed in a second pan containing culture water if additional organisms are needed for testing. The newborn amphipods are moved with a pipet and placed in a culture chamber with flowing water during a grow-out period. The newborn amphipods should be counted to determine if adequate numbers have been collected for the test.

10.3.4.3 Isolation of about 1500 (750 pairs) adults in amplexus provided about 800 newborn amphipods in 24 h and required about six man-hours of time. Isolation of about 4000 mixed-age adults (some in amplexus and others not in amplexus) provided about 800 newborn amphipods in 24 h and required less than one man-hour of time. The newborn amphipods should be held for 6 to 13 d to provide 7- to 14-d-old organisms to start a 10-d test (Section 11) or should be held for 7 d to provide 7- to 8-d-old organisms to start a long-term test (Section 14). The neonates are held in a 2-L beaker for 6 to 13 d before the start of a test. On the first day of isolation, the neonates are fed 10 mL of YCT (1800 mg/L stock solution) and 10 mL of *Selenastrum capricornutum* (about 3.5×10^7 cells/mL). On the third, fifth, seventh, ninth, eleventh, and thirteenth days after isolation, the amphipods are fed 5 mL of both YCT and *S. capricornutum*. Amphipods are initially fed a higher volume to establish a layer of food on the bottom of the culture chamber. If dissolved oxygen drops below 4 mg/L, about 50% of the water should be replaced (Ingersoll et al., 1998).

10.3.5 Laboratories that use mixed-age amphipods for testing must demonstrate that the procedure used to isolate amphipods will produce test organisms that are 7- to 14-d old. For example, amphipods passing through a U.S. Standard #35 sieve (500 µm), but stopped by a #45 sieve (355 µm) averaged 1.54 mm (SD 0.09) in length (P.V. Winger, USGS, Athens, GA, unpublished data). The mean length of these sieved organisms corresponds to that of 6-d-old amphipods (Figure 10.1). After holding for 3 d before testing to eliminate organisms injured during sieving, these amphipods would be about 9 d old (length 1.84 mm, SD 0.11) at the start of a toxicity test.

10.3.5.1 Ingersoll and Nelson (1990) describe the following procedure for obtaining mixed-age amphipods of a similar size to start a test. Smaller amphipods are isolated from larger amphipods using a stack of U.S. Standard sieves: #30 (600 µm), #40 (425 µm), and #60 (250 µm). Sieves should be held under water to isolate the amphipods. Amphipods may float on the surface of the water if they are exposed to air. Artificial substrate or leaves are placed in the #30 sieve. Culture water is rinsed through the sieves and small amphipods stopped by the #60 sieve are washed into a collecting pan. Larger amphipods in the #30 and #40 sieves are returned to the culture chamber. The smaller amphipods are then placed in 1-L beakers containing culture water and food (about 200 amphipods per beaker) with gentle aeration.

10.3.5.2 Amphipods should be held and fed at a rate similar to the mass cultures for at least 2 d before the start of a test to eliminate animals injured during handling.

10.3.6 See Section 10.2.6 for procedures used to evaluate the health of cultures.

10.4 Culturing Procedures for *Chironomus tentans*

10.4.1 The culturing methods described below are based on methods described in USEPA (1991a), Ankley et al. (1994a), Call et al. (1994), Greer (1993), ASTM (1999a), and USEPA (1994a). A *C. tentans* 10-d survival and growth test must be started with second- to third-instar larvae (about 10-d-old larvae; Section 12; Figure 10.2). At a temperature of 23°C, larvae should develop to the third instar by 9 to 11 d after hatching (about 11 to 13 d post-oviposition). The instar of midges at the start of a test can be determined based on head capsule width (Table 10.2) or based on weight or length at sediment test initiation. Average length of midge larvae should be 4 to 6 mm, while average dry weight should be 0.08 to 0.23 mg/individual. A *C. tentans* long-term test must be started with larvae less than 24 h old (see Section 15.3 for a

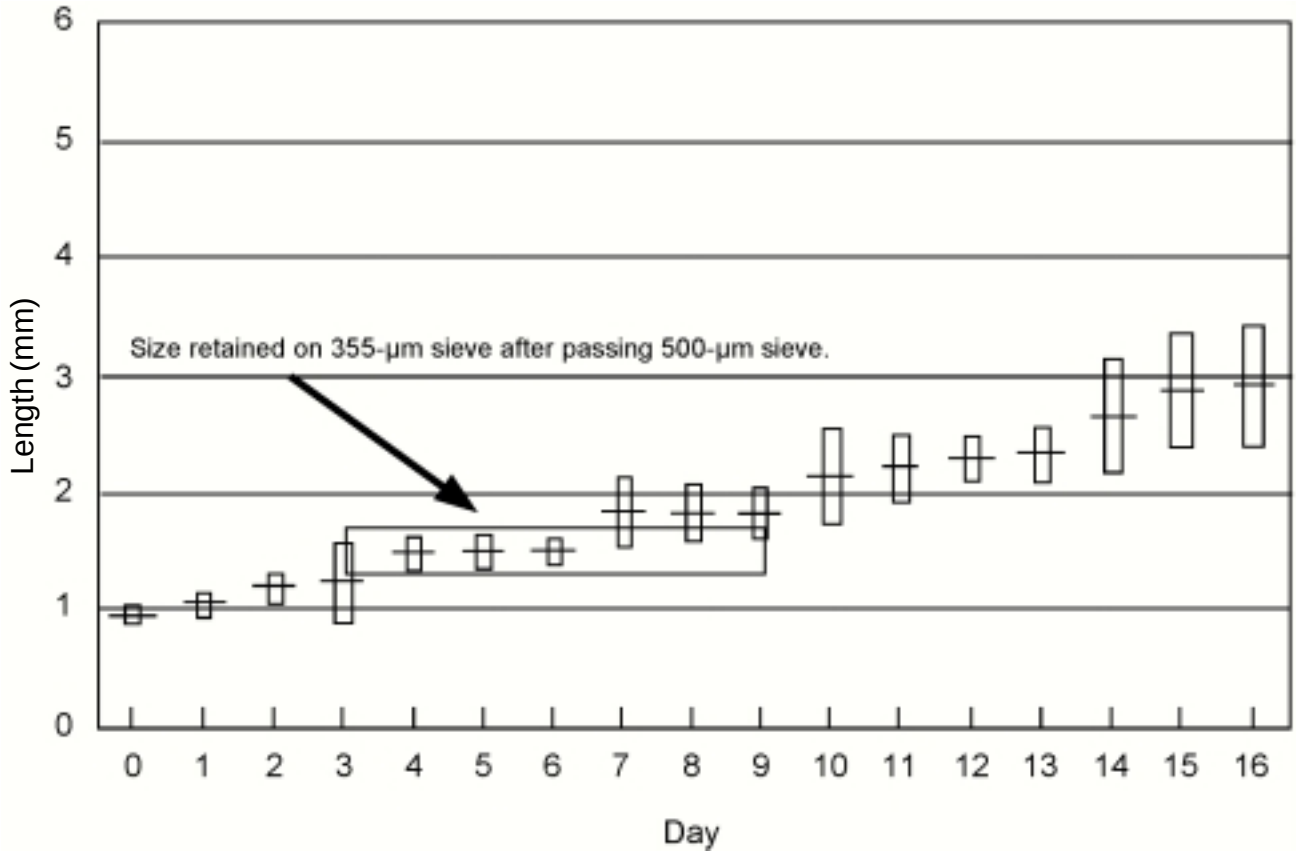


Figure 10.1 **Mean length (+/- 2SD) and relative age of *Hyalella azteca* collected by sieving in comparison with length of known-age organisms. P.V. Winger, USGS, Athens, GA, unpublished data.**

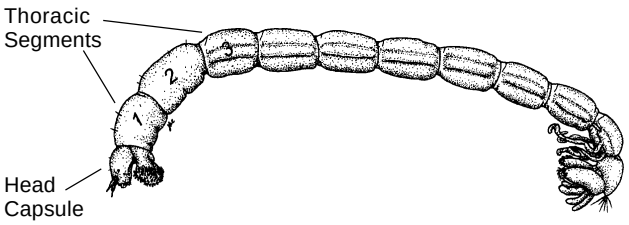


Figure 10.2. ***Chironomus tentans* larvae. Note thoracic segments which are used to measure instars. (Reprinted from Clifford, 1991 with kind permission from the University of Alberta Press.)**

Table 10.2 ***Chironomus tentans* Instar and Head Capsule Widths[1]**

Instar	Days after hatching	Mean (mm)	Range (mm)
First	1 to 4.4	0.10	0.09 to 0.13
Second	4.4 to 8.5	0.20	0.18 to 0.23
Third	8.5 to 12.5	0.38	0.33 to 0.45
Fourth	≥12.5	0.67	0.63 to 0.71

[1] T.J. Norberg-King, USEPA, Duluth, MN, unpublished data.

description of an approach for obtaining *C. tentans* larvae less than 24 h old).

10.4.2 Historically, third-instar *C. tentans* were frequently referred to as the second instar in the published literature. When *C. tentans* larvae were measured daily, the *C. tentans* raised at 22°C to 24°C were third instar, not second instar, by 9 to 11 d after hatching (T.J. Norberg-King, USEPA, Duluth, MN, unpublished data).

10.4.3 Both silica sand and shredded paper toweling have been used as substrates to culture *C. tentans.* Either substrate may be used if a healthy culture can be maintained. Greer (1993) used sand or paper toweling to culture midges; however, sand was preferred due to the ease in removing larvae for testing. Sources of sand are listed in Section 7.

10.4.3.1 Paper towels are prepared according to a procedure adapted from Batac-Catalan and White (1982). Plain white kitchen paper towels are cut into strips. Cut toweling is loosely packed into a blender with culture water and blended for a few seconds. Small pieces should be available to the organism; blending for too long will result in a fine pulp that will not settle in a culture tank. Blended towels can then be added directly to culture tanks, eliminating any conditioning period for the substrate. A mass of the toweling sufficient to fill a 150-mL beaker is placed into a blender containing 1 L of deionized water, and blended for 30 sec or until the strips are broken apart in the form of a pulp. The pulp is then sieved using a 710-µm

sieve and rinsed well with deionized water to remove the shortest fibers.

10.4.3.2 Dry shredded paper toweling loosely packed into a 2-L beaker will provide sufficient substrate for about ten 19-L chambers (USEPA, 1991a). The shredded toweling placed in a 150-mL beaker produces enough substrate for one 19-L chamber. Additional substrate can be frozen in deionized water for later use.

10.4.4 Five egg cases will provide a sufficient number of organisms to start a new culture chamber. Egg cases should be held at 23°C in a glass beaker or crystallizing dish containing about 100 to 150 mL of culture water (temperature change should not exceed 2°C per d). Food is not added until the embryos start to hatch (in about 2 to 4 d at 23°C) to reduce the risk of oxygen depletion. About 200 to 400 larvae are then placed into each culture chamber. Crowding of larvae will reduce growth. See Section 10.4.5.1 or 10.4.6.1 for a description of feeding rates. Larvae should reach the third instar by about 10 d after median hatch (about 12 to 14 d after the time the eggs were laid; Table 10.2).

10.4.5 *Chironomus tentans* are cultured in soft water at the USEPA laboratory in Duluth (USEPA, 1993c) in glass aquaria (19.0-L capacity, 36 x 21 x 26 cm high). A water volume of about 6 to 8 L in these flow-through chambers can be maintained by drilling an overflow hole in one end 11 cm from the bottom. The top of the aquarium is covered with a mesh material to trap emergent adults. Pantyhose with the elasticized waist is positioned around the chamber top and the legs are cut off. Fiberglass-window screen glued to a glass strip (about 2- to 3-cm wide) rectangle placed on top of each aquarium has also been used by Call et al. (1994). About 200 to 300 mL of 40-mesh silica sand is placed in each chamber.

10.4.5.1 The stocking density of the number of *C. tentans* eggs should be about 600 eggs per 6 to 8 L of water. Dawson et al. (1999) found that the cultures in 15-L aquaria and 7 L of water were self-regulating in density regardless of the initial number of eggs stocked in each tank. However, tanks with a higher initial stocking density (i.e., 1400 eggs/tank) increased the time of peak adult emergence to 30 to 33 d, whereas tanks with lower stocking densities (600 or 1000 eggs/tank) had peak emergence at 22 to 25 d after hatching.

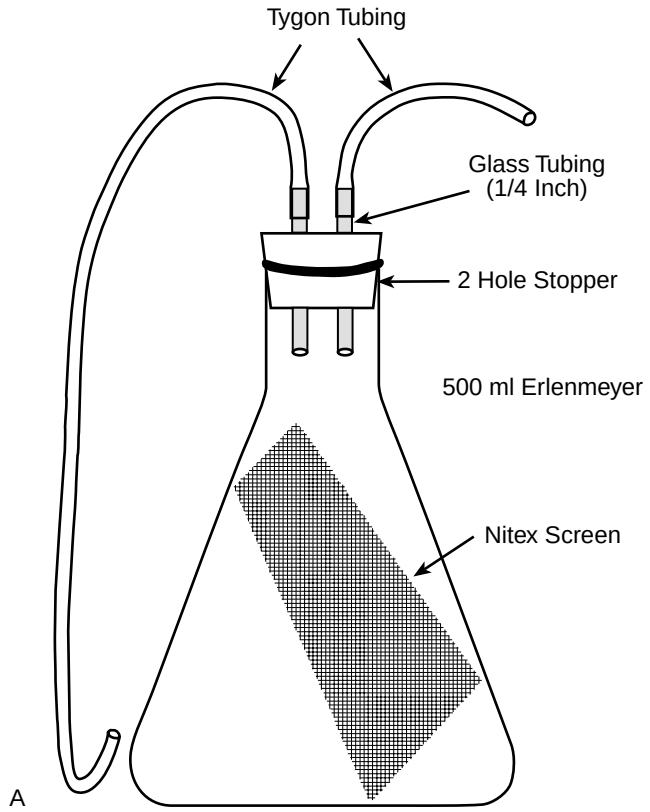
10.4.5.2 Fish food flakes (i.e., Tetrafin®) are added to each culture chamber to provide a final food concentration of about 0.04 mg dry solids/mL of culture water. A stock suspension of the solids is blended with distilled water to form an initial slurry. It is then filtered through a 200-micron Nitex screen and diluted with distilled water to form a 56 g dry solids/L final slurry (Appendix B). The larvae in each tank are fed 2.5 mL of slurry (140 mg of Tetrafin per day) from Day 0 to Day 7 and 5 mL of slurry (280 mg Tetrafin per day) from Day 8 on. Feeding is done after the water renewal process is completed. The stock suspension should be well mixed immediately before removing an aliquot for feeding. Each batch of food should be refrigerated and can be used for up to two weeks (Appendix B). Laboratories using static systems should develop lower feeding rates specific to their systems.

10.4.6 *Chironomus tentans* are cultured by Greer (1993) in Rubbermaid® 5.7-L polyethylene cylindrical containers. The containers are modified by cutting a semicircle into the lid 17.75 cm across by 12.5 cm. Stainless-steel screen (20 mesh/0.4 cm) is cut to size and melted to the plastic lid. The screen provides air exchange, retains emerging adults, and is a convenient way to observe the culture. Two holes about 0.05 cm in diameter are drilled through the uncut portion of the lid to provide access for an air line and to introduce food. The food access hole is closed with a No. 00 stopper. Greer (1993) cultures midges under static conditions with moderate aeration, and about 90% of the water is replaced weekly. Each 5.7-L culture chamber contains about 3 L of water and about 25 mL of fine sand. Eight to 10 chambers are used to maintain the culture.

10.4.6.1 Midges in each chamber are fed 6 mL/d of a 100 g/L suspension of fish food flakes (e.g., Tetrafin®) on Tuesday, Wednesday, Thursday, Friday, and Sunday. A 6-mL chlorella suspension (deactivated "Algae-Feast® *Chlorella,*" Earthrise Co., Callpatria, CA) is added to each chamber on Saturday and on Monday. The chlorella suspension is prepared by adding 5 g of dry chlorella powder/L of water. The mixture should be refrigerated and can be used for up to two weeks.

10.4.6.2 The water should be replaced more often if animals appear stressed (e.g., at surface or pale color at the second instar) or if the water is cloudy. Water is replaced by first removing emergent adults with an aspirator. Any growth on the sides of the chamber should be brushed off before water is removed. Care should be taken not to pour or siphon out the larvae when removing the water. Larvae will typically stay near the bottom; however, a small-mesh sieve or nylon net can be used to catch any larvae that float out. After the chambers have been cleaned, temperature-adjusted culture water is poured back into each chamber. The water should be added quickly to stir up the larvae. Using this procedure, the approximate size, number, and the general health of the culture can be observed.

10.4.7 Adult emergence will begin about three weeks after hatching at 23°C. Once adults begin to emerge, they can be gently siphoned into a dry aspirator flask on a daily basis. An aspirator can be made using a 250- or 500-mL Erlenmeyer flask, a two-hole stopper, some short sections of 0.25-inch glass tubing, and Tygon® tubing for collecting and providing suction (Figure 10.3). Adults should be aspirated with short inhalations to avoid injuring the organisms. The mouthpiece on the aspirator should be replaced or disinfected between use. Sex ratio of the adults should be checked to ensure that a sufficient number of males are available for mating and fertilization. One male may fertilize more than one female. However, a



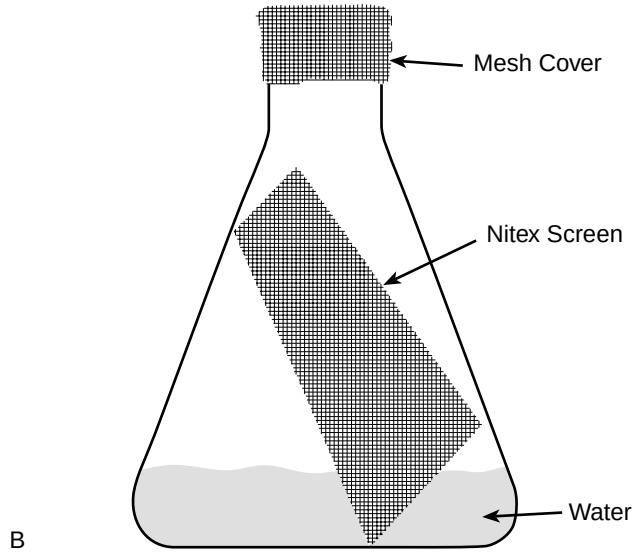


Figure 10.3 Aspirator chamber (A) and reproduction and oviposit chamber (B) for adult midges.

ratio of one male to three females improves fertilization success.

10.4.7.1 A reproduction and oviposit chamber may be prepared in several different ways (Figure 10.3). Culture water (about 50 to 75 mL) can be added to the aspiration flask in which the adults were collected (Figure 10.3; Batac-Catalan and White, 1982). The USEPA Office of Research and Development Laboratory (Duluth, MN; USEPA, 1991a) uses a 500-mL collecting flask with a length of Nitex® screen positioned vertically and extending into the culture water (Figure 10.3). The Nitex® screen is used by the females to position themselves just above the water during oviposition. The two-hole stopper and tubing of the aspirator should be replaced by screened material or a cotton plug for good air exchange in the oviposition chamber.

10.4.7.2 Greer (1993) uses an oviposition box to hold emergent adults. The box is constructed of a 5.7-L chamber with a 20-cm tall cylindrical chamber on top. The top chamber is constructed of stainless-steel screen (35 mesh/ 2.54 cm) melted onto a plastic lid with a 17.75-cm hole. A 5-cm hole is cut into the side of the bottom chamber and a #11 stopper is used to close the hole. Egg cases are removed by first sliding a piece of plexiglass between the top and bottom chambers. Adult midges are then aspirated from the bottom chamber. The top chamber with plexiglass is removed from the bottom chamber and a forceps is used to remove the egg cases. The top chamber is put back on top of the bottom chamber, the plexiglass is removed, and the aspirated adults are released from the aspirator into the chamber through the 5-cm hole.

10.4.8 About two to three weeks before the start of a test, at least 3 to 5 egg cases should be isolated for hatching using procedures outlined in Section 10.4.4.

10.4.9 Records should be kept on the time to first emergence and the success of emergence for each culture chamber. It is also desirable to monitor growth and head capsule width periodically in the cultures. See Section 10.2.6 for additional detail on procedures for evaluating the health of the cultures.

10.5 Culturing Procedures for *Lumbriculus variegatus*

10.5.1 The culturing procedures described below are based on methods described in Phipps et al. (1993), USEPA (1991a), Call et al. (1994), Brunson et al. (1998), and USEPA (1994a). Bioaccumulation tests are started with adult organisms.

10.5.2 *Lumbriculus variegatus* are generally cultured with daily renewal of water (57- to 80-L aquaria containing 45 to 50 L of water).

10.5.3 Paper towels can be used as a substrate for culturing *L. variegatus* (Phipps et al., 1993). Substrate is prepared by cutting unbleached brown paper towels into strips either with a paper shredder or with scissors. Cut toweling is loosely packed into a blender with culture water and blended for a few seconds. Small pieces should be available to the organisms; blending for too long will result in a fine pulp that will not settle in culture tanks. Blended towels can then be added directly to culture tanks, eliminating any conditioning period for the substrate. The paper towel substrate is renewed with blended towels when thin or bare areas appear in the cultures. The substrate in the chamber will generally last for about two months.

10.5.4 Oligochaetes probably obtain nourishment from ingesting the organic matter in the substrate (Pennak, 1989). *Lumbriculus variegatus* in each of the culture chambers are fed a 10-mL suspension of 6 g of trout starter 3 times/week. The particles will temporarily disperse on the surface film, break through the surface tension, and settle out over the substrate. Laboratories using static systems should develop lower feeding rates specific to their systems. Food and substrate used to culture oligochaetes should be analyzed for compounds to be evaluated in bioaccumulation tests. If the concentration of the test compound is above the detection level and the food is not measured, the test may be invalidated. Recent studies in other laboratories, for example, have indicated elevated concentrations of PCBs in substrate and/or food used for culturing the oligochaete (J. Amato, AScI Corporation, Duluth, MN, personal communication).

10.5.5 Phipps et al. (1993) recommend starting a new culture with 500 to 1000 worms. Conditioned paper toweling should be added when the substrate in a culture chamber is thin.

10.5.6 On the day before the start of a test, oligochaetes can be isolated by transferring substrate from the cultures into a beaker using a fine-mesh net. Additional organisms can be removed using a glass pipet (20-cm long, 5-mm i.d.; Phipps et al., 1993). Water can be slowly trickled into the beaker. The oligochaetes will form a mass and most of the remaining substrate will be flushed from the beaker. On the day the test is started, organisms can be placed in glass or stainless-steel pans. A gentle stream of water from the pipet can be used to spread out clusters of oligochaetes. The remaining substrate can be siphoned from the pan by allowing the worms to reform in a cluster on the bottom of the pan. For bioaccumulation tests, aliquots of worms to be added to each test chamber can be transferred using a blunt dissecting needle or dental pick. Excess water can be removed during transfer by touching the mass of oligochaetes to the edge of the pan. The mass of oligochaetes is then placed in a tared weigh boat, quickly weighed, and immediately introduced into the appropriate test chamber. Organisms should not be blotted with a paper towel to remove excess water (Brunson et al., 1998).

10.5.7 The culture population generally doubles (number of organisms) in about 10 to 14 d. See Section 10.2.6 for additional detail on procedures for evaluating the health of the cultures.

Section 11
Test Method 100.1
Hyalella azteca 10-d Survival and Growth Test for Sediments

11.1 Introduction

11.1.1 *Hyalella azteca* (Saussure) have many desirable characteristics of an ideal sediment toxicity testing organism including relative sensitivity to contaminants associated with sediment, short generation time, contact with sediment, ease of culture in the laboratory, and tolerance to varying physico-chemical characteristics of sediment. Their response has been evaluated in interlaboratory studies and has been confirmed with natural benthic populations. Many investigators have successfully used *H. azteca* to evaluate the toxicity of freshwater sediments (e.g., Nebeker et al., 1984a; Borgmann and Munwar, 1989; Ingersoll and Nelson, 1990; Ankley et al., 1991a; Ankley et al., 1991b; Burton et al., 1989; Winger and Lasier, 1993; Kemble et al., 1994). *H. azteca* has been used for a variety of sediment assessments (Ankley et al., 1991; West et al., 1993; Hoke et al., 1994, 1995; West et al., 1994). *Hyalella azteca* can also be used to evaluate the toxicity of estuarine sediments (up to 15 ‰ salinity; Nebeker and Miller, 1988; Roach et al., 1992; Winger et al., 1993). Endpoints typically monitored in 10-d sediment toxicity tests with *H. azteca* include survival and growth.

11.1.2 A test method for conducting a 10-d sediment toxicity test is described in Section 11.2 for *H. azteca.* Methods outlined in Appendix A of USEPA (1994a) and in Section 11.1.1 were used for developing test method 100.1. Results of tests using procedures different from the procedures described in Section 11.2 may not be comparable, and these different procedures may alter contaminant bioavailability. Comparison of results obtained using modified versions of these procedures might provide useful information concerning new concepts and procedures for conducting sediment tests with aquatic organisms. If tests are conducted with procedures different from the procedures described in this manual, additional tests are required to determine comparability of results (Section 1.3).

11.2 Recommended Test Method for Conducting a 10-d Sediment Toxicity Test with *Hyalella azteca*

11.2.1 Recommended conditions for conducting a 10-d sediment toxicity test with *H. azteca* are summarized in Table 11.1. A general activity schedule is outlined in Table 11.2. Decisions concerning the various aspects of experimental design, such as the number of treatments, number of test chambers/treatment, and water-quality characteristics should be based on the purpose of the test and the methods of data analysis (Section 16). The number of replicates and concentrations tested depends in part on the significance level selected and the type of statistical analysis. When variability remains constant, the sensitivity of a test increases as the number of replicates increase.

11.2.2 The recommended 10-d sediment toxicity test with *H. azteca* must be conducted at 23°C with a 16L:8D photoperiod at an illuminance of about 100 to 1000 lux (Table 11.1). Test chambers are 300-mL high-form lipless beakers containing 100 mL of sediment and 175 mL of overlying water. Ten 7- to 14-d-old amphipods are used to start a test. The 10-d test should start with a narrow range in size or age of *H. azteca* (i.e., 1- to 2-d range in age) to reduce potential variability in growth at the end of a 10-d test (Section 10.3.1). The number of replicates/treatment depends on the objective of the test. Eight replicates are recommended for routine testing (Section 16). Amphipods in each test chamber are fed 1.0 mL of YCT food daily (Appendix B). The first edition of the manual (USEPA, 1994a) recommended a feeding level of 1.5 mL of YCT daily; however, this feeding level was revised to 1.0 mL to be consistent, with the feeding level in the long-term test with *H. azteca* (Section 14). Each chamber receives 2 volume additions/d of overlying water. Water renewals may be manual or automated. Appendix A describes water-renewal systems that can be used to deliver overlying water. Overlying water can be culture water, well water, surface water, site water, or reconstituted water. For site-specific evaluations, the characteristics of the overlying water should be as similar as possible to the site where sediment is collected. Requirements for test acceptability are summarized in Table 11.3.

11.3 General Procedures

11.3.1 Sediment into Test Chambers

11.3.1.1 The day before the sediment test is started (Day -1) each sediment should be thoroughly homogenized and added to the test chambers (Section 8.3.1). Sediment should be visually inspected to judge the degree of homogeneity. Excess water on the surface of the sediment can indicate separation of solid and liquid components. If a quantitative measure of homogeneity is

Table 11.1 Test Conditions for Conducting a 10-d Sediment Toxicity Test with *Hyalella azteca*

	Parameter	Conditions
1.	Test type:	Whole-sediment toxicity test with renewal of overlying water
2.	Temperature:	23 ± 1°C
3.	Light quality:	Wide-spectrum fluorescent lights
4.	Illuminance:	About 100 to 1000 lux
5.	Photoperiod:	16L:8D
6.	Test chamber:	300-mL high-form lipless beaker
7.	Sediment volume:	100 mL
8.	Overlying water volume:	175 mL
9.	Renewal of overlying water:	2 volume additions/d (Appendix A); continuous or intermittent (e.g., 1 volume addition every 12 h)
10.	Age of organisms:	7- to 14-d old at the start of the test (1- to 2-d range in age)
11.	Number of organisms/chamber:	10
12.	Number of replicate chambers/treatment:	Depends on the objective of the test. Eight replicates are recommended for routine testing (see Section 16).
13.	Feeding:	YCT food, fed 1.0 mL daily (1800 mg/L stock) to each test chamber. The first edition of the manual (USEPA, 1994a) recommended a feeding level of 1.5 mL of YCT daily; however, this feeding level was revised to 1.0 mL to be consistent with the feeding level in the long-term tests with *H. azteca* (Section 14).
14.	Aeration:	None, unless dissolved oxygen in overlying water drops below 2.5 mg/L.
15.	Overlying water:	Culture water, well water, surface water, site water, or reconstituted water
16.	Test chamber cleaning:	If screens become clogged during a test, gently brush the *outside* of the screen (Appendix A).
17.	Overlying water quality:	Hardness, alkalinity, conductivity, pH, and ammonia at the beginning and end of a test. Temperature and dissolved oxygen daily.
18.	Test duration:	10 d
19.	Endpoints:	Survival and growth
20.	Test acceptability:	Minimum mean control survival of 80% and measurable growth of test organisms in the control sediment. Additional performance-based criteria specifications are outlined in Table 11.3.

Table 11.2 General Activity Schedule for Conducting a 10-d Sediment Toxicity Test with *Hyalella azteca* [1]

Day	Activity
-7	Separate known-age amphipods from the cultures and place in holding chambers. Begin preparing food for the test. There should be a 1- to 2-d range in age of amphipods used to start the test.
-6 to -2	Feed and observe isolated amphipods (Section 10.3), monitor water quality (e.g., temperature and dissolved oxygen).
-1	Feed and observe isolated amphipods (Section 10.3), monitor water quality. Add sediment into each test chamber, place chambers into exposure system, and start renewing overlying water.
0	Measure total water quality (pH, temperature, dissolved oxygen, hardness, alkalinity, conductivity, ammonia). Transfer 10 7- to 14-day-old amphipods into each test chamber. Release organisms under the surface of the water. Add 1.0 mL of YCT into each test chamber. Archive 20 test organisms for length determination or archive 80 test organisms for dry weight determination. Observe behavior of test organisms.
1 to 8	Add 1.0 mL of YCT food to each test chamber. Measure temperature and dissolved oxygen. Observe behavior of test organisms.
9	Measure total water quality.
10	Measure temperature and dissolved oxygen. End the test by collecting the amphipods with a sieve (Section 11.3.7.1). Count survivors and prepare organisms for weight or length measurements.

[1] Modified from Call et al., 1994

Table 11.3 Test Acceptability Requirements for a 10-d Sediment Toxicity Test with *Hyalella azteca*

A. It is recommended for conducting a 10-d test with *Hyalella azteca* that the following performance criteria be met:

1. Age of *H. azteca* at the start of the test must be between 7- to 14-d old. The 10-d test should start with a narrow range in size or age of *H. azteca* (i.e., 1- to 2-d range in age) to reduce potential variability in growth at the end of a 10-d test (Section 10.3.1).
2. Average survival of *H. azteca* in the control sediment must be greater than or equal to 80% at the end of the test. Growth of test organisms should be measurable in the control sediment at the end of the 10-d test (i.e., relative to organisms at the start of the test).
3. Hardness, alkalinity, and ammonia in the overlying water typically should not vary by more than 50% during the test, and dissolved oxygen should be maintained above 2.5 mg/L in the overlying water.

B. Performance-based criteria for culturing *H. azteca* include the following:

1. It may be desirable for laboratories to periodically perform 96-h water-only reference-toxicity tests to assess the sensitivity of culture organisms (Section 9.16.2). Data from these reference-toxicity tests could be used to assess genetic strain or life-stage sensitivity of test organisms to select chemicals.
2. Laboratories should track parental survival in the cultures and record this information using control charts if known-age cultures are maintained. Records should also be kept on the frequency of restarting cultures and the age of brood organisms.
3. Laboratories should record the following water-quality characteristics of the cultures at least quarterly: pH, hardness, alkalinity, and ammonia. Dissolved oxygen in the cultures should be measured weekly. Temperature of the cultures should be recorded daily. If static cultures are used, it may be desirable to measure water quality more frequently.
4. Laboratories should characterize and monitor background contamination and nutrient quality of food if problems are observed in culturing or testing organisms.
5. Physiological measurements such as lipid content might provide useful information regarding the health of the cultures.

C. Additional requirements:

1. All organisms in a test must be from the same source.
2. Storage of sediments collected from the field should follow guidance outlined in Section 8.2.
3. All test chambers (and compartments) should be identical and should contain the same amount of sediment and overlying water.
4. Negative-control sediment and appropriate solvent controls must be included in a test. The concentration of solvent used must not adversely affect test organisms.
5. Test organisms must be cultured and tested at 23°C (±1°C).
6. The daily mean test temperature must be within ±1°C of 23°C. The instantaneous temperature must always be within ±3°C of 23°C.
7. Natural physico-chemical characteristics of test sediment collected from the field should be within the tolerance limits of the test organisms.

required, replicate subsamples should be taken from the sediment batch and analyzed for TOC, chemical concentrations, and particle size.

11.3.1.2 Each test chamber should contain the same amount of sediment, determined either by volume or by weight. Overlying water is added to the chambers on Day -1 in a manner that minimizes suspension of sediment. This can be accomplished by gently pouring water along the sides of the chambers or by pouring water onto a baffle (e.g., a circular piece of Teflon® with a handle attached) placed above the sediment to dissipate the force of the water. A test begins when the organisms are added to the test chambers (Day 0).

11.3.2 Renewal of Overlying Water

11.3.2.1 Renewal of overlying water is required during a test. At any particular time during the test, flow rates through any two test chambers should not differ by more than 10%. Hardness, alkalinity and ammonia concentrations in the water above the sediment, within a treatment, typically should not vary by more than 50% during the test. Mount and Brungs (1967) diluters have been modified for sediment testing, and other automated water-delivery systems have also been used (Maki, 1977; Ingersoll and Nelson, 1990; Benoit et al., 1993; Zumwalt et al., 1994; Brunson et al., 1998; Wall et al., 1998; Leppanen and Maier, 1998). The water-delivery system should be calibrated before a test is started to verify that the system is functioning properly. Renewal of overlying water is started on Day -1 before the addition of test organisms or food on Day 0. Appendix A describes water-renewal systems that can be used for conducting sediment tests.

11.3.2.2 In water-renewal tests with one to four volume additions of overlying water/d, water-quality characteristics generally remain similar to the inflowing water (Ingersoll and Nelson, 1990; Ankley et al., 1993); however, in static

tests, water quality may change profoundly during the exposure (Shuba et al., 1978). For example, in static whole-sediment tests, the alkalinity, hardness, and conductivity of overlying water more than doubled in several treatments during a four-week exposure (Ingersoll and Nelson, 1990). Additionally, concentrations of metabolic products (e.g., ammonia) may also increase during static exposures, and these compounds can either be directly toxic to the test organisms or may contribute to the toxicity of the contaminants in the sediment. Furthermore, changes in water-quality characteristics such as hardness may influence the toxicity of many inorganic (Gauss et al., 1985) and organic (Mayer and Ellersieck, 1986) contaminants. Although contaminant concentrations are reduced in the overlying water in water-renewal tests, organisms in direct contact with sediment generally receive a substantial proportion of a contaminant dose directly from either the whole sediment or from the pore water.

11.3.3 Acclimation

11.3.3.1 Test organisms must be cultured and tested at 23°C. Ideally, test organisms should be cultured in the same water that will be used in testing. However, acclimation of test organisms to the test water is not required.

11.3.3.2 Culturing of organisms and toxicity assessment are typically conducted at 23°C. However, occasionally there is a need to perform evaluations at temperatures different than that recommended. Under these circumstances, it may be necessary to acclimate organisms to the desired test temperature to prevent thermal shock when moving immediately from the culture temperature to the test temperature (ASTM, 1999a). Acclimation can be achieved by exposing organisms to a gradual change in temperature; however, the rate of change should be relatively slow to prevent thermal shock. A change in temperature of 1°C every 1 to 2 h has been used successfully in some studies (P.K. Sibley, University of Guleph, Guelph, Ontario, personal communication; APHA, 1989). Testing at temperatures other than 23°C needs to be preceded by studies to determine expected performance under alternate conditions.

11.3.4 Placing Organisms in Test Chambers

11.3.4.1 Test organisms should be handled as little as possible. Amphipods should be introduced into the overlying water below the air-water interface. Test organisms can be pipetted directly into overlying water. The size of the test organisms at the start of the test should be measured using the same measure (length or weight) that will be used to assess their size at the end of the test. For length, a minimum of 20 organisms should be measured. For weight measurement, a larger sample size (e.g., 80) may be desirable because of the relative small mass of the organisms. This information can be used to determine consistency in the size of the organisms used to start a test.

11.3.5 Feeding

11.3.5.1 For each beaker, 1.0 mL of YCT is added from Day 0 to Day 9. Without addition of food, the test organisms may starve during exposures. However, the addition of the food may alter the availability of the contaminants in the sediment (Wiederholm et al., 1987; Harkey et al., 1994). Furthermore, if too much food is added to the test chamber or if the mortality of test organisms is high, fungal or bacterial growth may develop on the sediment surface. Therefore, the amount of food added to the test chambers is kept to a minimum.

11.3.5.2 Suspensions of food should be thoroughly mixed before aliquots are taken. If excess food collects on the sediment, a fungal or bacterial growth may develop on the sediment surface, in which case feeding should be suspended for one or more days. A drop in dissolved oxygen below 2.5 mg/L during a test may indicate that the food added is not being consumed. Feeding should be suspended for the amount of time necessary to increase the dissolved oxygen concentration (ASTM, 1999a). If feeding is suspended in one treatment, it should be suspended in all treatments. Detailed records of feeding rates and the appearance of the sediment surface should be made daily.

11.3.6 Monitoring a Test

11.3.6.1 All chambers should be checked daily and observations made to assess test organism behavior such as sediment avoidance. However, monitoring effects on burrowing activity of test organisms may be difficult because the test organisms are often not visible during the exposure. The operation of the exposure system should be monitored daily.

11.3.6.2 Measurement of Overlying Water-quality Characteristics

11.3.6.2.1 Conductivity, hardness, pH, alkalinity, and ammonia should be measured in all treatments at the beginning and end of a test. Overlying water should be sampled just before water renewal from about 1 to 2 cm above the sediment surface using a pipet. It may be necessary to composite water samples from individual replicates. The pipet should be checked to make sure no organisms are removed during sampling of overlying water. Water quality should be measured on each batch of water prepared for the test.

11.3.6.2.2 Dissolved oxygen should be measured daily and should be maintained at a minimum of 2.5 mg/L. If a probe is used to measure dissolved oxygen in overlying water, it should be thoroughly inspected between samples to make sure that organisms are not attached and should be rinsed between samples to minimize cross contamination. Aeration can be used to maintain dissolved oxygen in the overlying water above 2.5 mg/L (i.e., about 1 bubble/second in the overlying water). Dissolved oxygen and pH can be measured directly in the overlying water with a probe.

11.3.6.2.3 Temperature should be measured at least daily in at least one test chamber from each treatment. The temperature of the water bath or the exposure chamber should be continuously monitored. The daily mean test temperature must be within ±1°C of 23°C. The instantaneous temperature must always be within ±3°C of 23°C.

11.3.7 Ending a Test

11.3.7.1 Any of the surviving amphipods in the water column or on the surface of the sediment can be pipetted from the beaker before sieving the sediment. Immobile organisms isolated from the sediment surface or from sieved material should be considered dead. A #40 sieve (425-µm mesh) can be used to remove amphipods from sediment. Alternatively, Kemble et al. (1994) suggest sieving of sediment using the following procedure: (1) pour about half of the overlying water through a #50- (300-µm) U.S. standard mesh sieve, (2) swirl the remaining water to suspend the upper 1 cm of sediment, (3) pour this slurry through the #50-mesh sieve and wash the contents of the sieve into an examination pan, (4) rinse the coarser sediment remaining in the test chamber through a #40- (425-µm) mesh sieve and wash the contents of this second sieve into a second examination pan. Surviving test organisms are removed from the two pans and counted. If growth (length) is to be measured (Ingersoll and Nelson, 1990), the organisms can be preserved in 8% sugar formalin solution. The sugar formalin solution is prepared by adding 120 g of sucrose to 80 mL of formalin, which is then brought to a volume of 1 L using deionized water. This stock solution is mixed with an equal volume of deionized water when used to preserve organisms. NoTox® (Earth Safe Industries, Belle Mead, NJ) can be used as a substitute for formalin (Unger et al., 1993).

11.3.7.2 A consistent amount of time should be taken to examine sieved material for recovery of test organisms (e.g., 5 min/replicate). Laboratories should demonstrate that their personnel are able to recover an average of at least 90% of the organisms from whole sediment. For example, test organisms could be added to control or test sediments, and recovery could be determined after 1 h (Tomasovic et al., 1994).

11.3.8 Test Data

11.3.8.1 Survival and growth are measured at the end of the 10-d sediment toxicity test with *H. azteca.* Growth of amphipods is often a more sensitive toxicity endpoint compared to survival (Burton and Ingersoll, 1994; Kemble et al., 1994; Becker et al., 1995; Ingersoll et al., 1996; Ingersoll et al., 1998; Steevens and Benson, 1998). The duration of the 10-d test starting with 7- to 14-d-old amphipods is not long enough to determine sexual maturation or reproductive effects. The 42-d test (Section 14) is designed to evaluate additional sublethal endpoints in sediment toxicity tests with *H. azteca*. See Section 14.4.5.3 for a discussion of measuring dry weight vs. length of *H. azteca*.

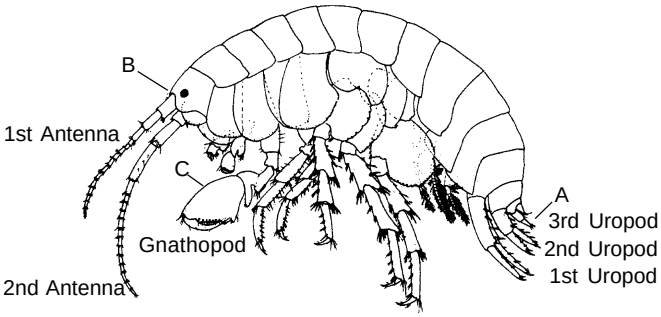


Figure 11.1 ***Hyalella azteca*. (A) denotes the uropods; (B) denotes the base of the first antennae; (C) denotes the gnathopod used for grasping females. Meaurement of length is made from base of the 3rd uropod (A) to (B). Females are recognized by the presence of egg cases or the absence of an enlarged gnathopod. (Reprinted from Cole and Watkins, 1997 with kind permission from Kluwer Academic Publishers.)**

11.3.8.2 Amphipod body length (±0.1 mm) can be measured from the base of the first antenna to the tip of the third uropod along the curve of the dorsal surface (Figure 11.1). Ingersoll and Nelson (1990) describe the use of a digitizing system and microscope to measure lengths of *H. azteca.* Kemble et al. (1994) also photographed invertebrates (at a magnification of 3.5X) and measured length using a computer-interfaced digitizing tablet. Antennal segment number can also be used to estimate length or weight of amphipods (E.L. Brunson, USGS, Columbia, MO, personal communication). Wet or dry weight measurements have also been used to estimate growth of *H. azteca* (ASTM, 1999a). If test organisms are to be used for an evaluation of bioaccumulation, it is not advisable to dry the sample before conducting the residue analysis. If conversion from wet weight to dry weight is necessary, aliquots of organisms can be weighed to establish wet to dry weight conversion factors. A consistent procedure should be used to remove the excess water from the organisms before measuring wet weight.

11.3.8.3 Dry weight of amphipods should be determined by pooling all living organisms from a replicate and drying the sample at about 60°C to 90°C to a constant weight. The sample is brought to room temperature in a desiccator and weighed to the nearest 0.01 mg to obtain mean weight per surviving organism per replicate (see Section 14.3.7.6) The first edition of this manual (USEPA, 1994a) recommended dry weight as a measure of growth for both *H. azteca* and *C. tentans*. For *C. tentans*, this recommendation was changed in the current edition to ash-free dry weight (AFDW) instead of dry weight, with the intent of reducing bias introduced by gut contents (Sibley et al., 1997a). However, this recommendation was not extended to include *H. azteca*. Studies by Dawson et al. (personal communication, T.D. Dawson, Integrated Laboratory Systems, Duluth, MN) have indicated that the ash content of *H. azteca* is not greatly decreased by purging organisms in clean water before weighing, suggesting that sediment does not comprise a large portion of the overall dry weight. In addition, using AFDW further decreases an

already small mass, potentially increasing measurement error. For this reason, dry weight continues to be the recommended endpoint for estimating growth of *H. azteca* via weight (growth can also be determined via length).

11.4 Interpretation of Results

11.4.1 Section 16 describes general information for interpretation of test results. The following sections describe species-specific information that is useful in helping to interpret the results of sediment toxicity tests with *H. azteca.*

11.4.2 Age Sensitivity

11.4.2.1 The sensitivity of *H. azteca* appears to be relatively similar up to at least 24- to 26-d-old organisms (Collyard et al., 1994). For example, the toxicity of diazinon, Cu, Cd, and Zn was similar in 96-h water-only exposures starting with 0- to 2-d-old organisms through 24- to 26-d-old organisms (Figure 11.2). The toxicity of alkylphenol ethoxylate (a surfactant) tended to increase with age. In general, this suggests that tests started with 7- to 14-d-old amphipods would be representative of the sensitivity of *H. azteca* up to at least the adult life stage.

11.4.3 Grain Size

11.4.3.1 *Hyalella azteca* are tolerant of a wide range of substrates. Physico-chemical characteristics (e.g., grain size or TOC) of sediment were not significantly correlated to the response of *H. azteca* in toxicity tests in which organisms were fed (Section 10.1.1.8; Ankley et al., 1994a).

11.4.4 Isolating Organisms at the End of a Test

11.4.4.1 Quantitative recovery of young amphipods (e.g., 0- to 7-d old) is difficult given their small size (Figure 11.3, Tomasovic et al., 1994). Recovery of older and larger amphipods (e.g., 21-d old) is much easier. This was a primary reason for deciding to start 10-d tests with 7- to 14-d-old amphipods (organisms are 17- to 24-d old at the end of the 10-d test).

11.4.5 Influence of Indigenous Organisms

11.4.5.1 Survival of *H. azteca* in 28-d tests was not reduced in the presence of oligochaetes in sediment samples (Reynoldson et al., 1994). However, growth of amphipods was reduced when high numbers of oligochaetes were placed in a sample. Therefore, it is important to determine the number and biomass of indigenous organisms in field-collected sediment in order to better interpret growth data (Reynoldson et al., 1994; DeFoe and Ankley, 1998). Furthermore, presence of predators may also influence the response of test organisms in sediment (Ingersoll and Nelson, 1990).

11.4.6 Ammonia toxicity

11.4.6.1 Section 1.3.7.5 addresses interpretative guidance for evaluating toxicity associated with ammonia in sediment.

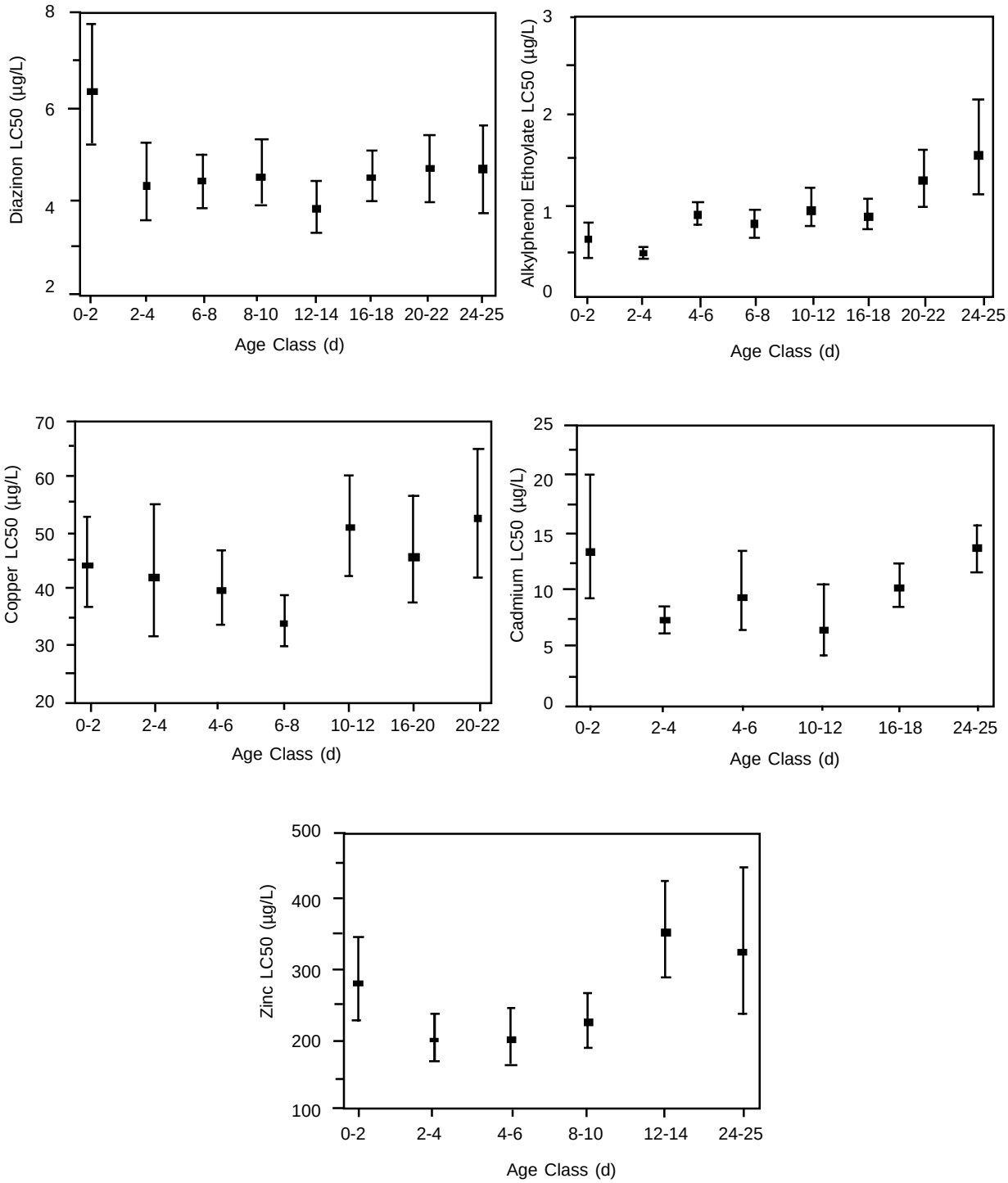


Figure 11.2 Lifestage sensitivity of *Hyalella azteca* in 96-h water-only exposures.

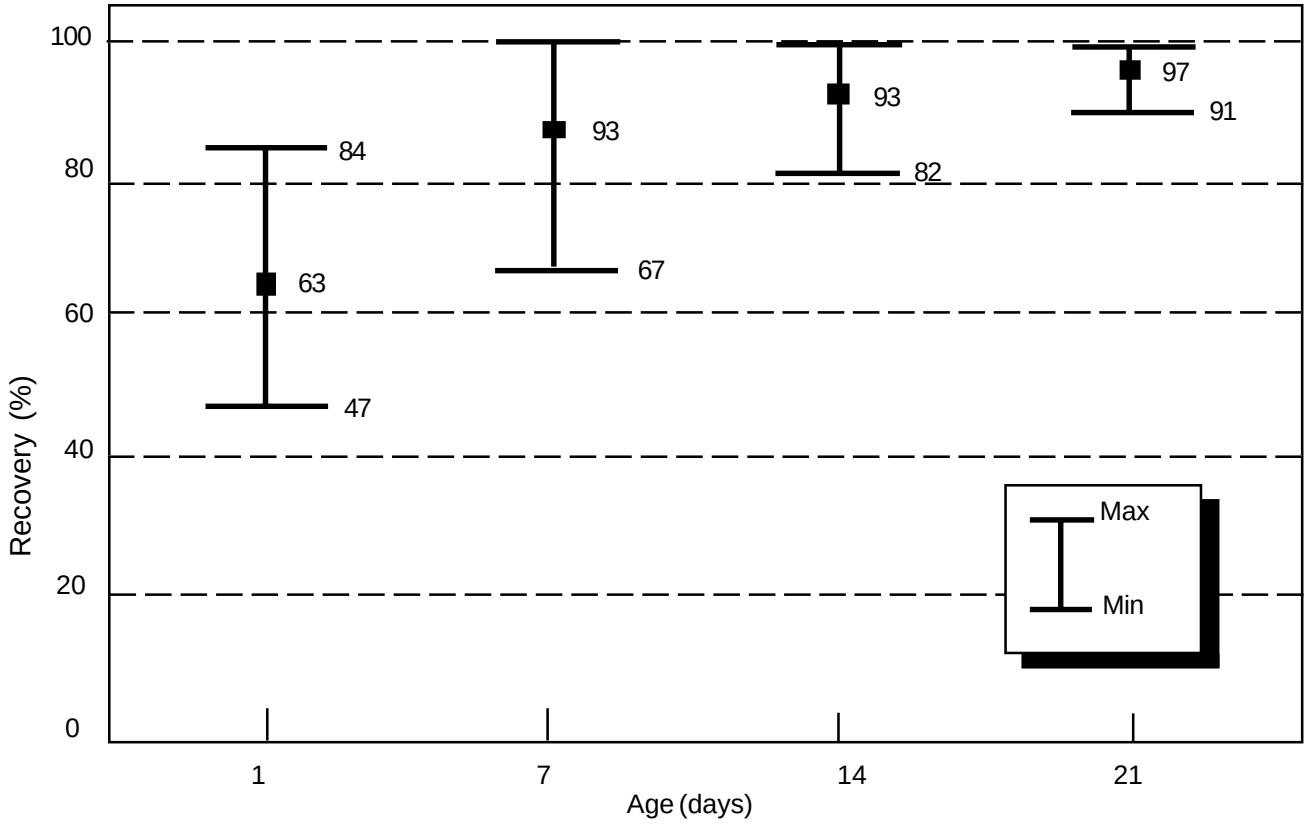


Figure 11.3 Average recovery of different age *Hyalella azteca* from sediment by 7 individuals.

Section 12
Test Method 100.2
Chironomus tentans 10-d Survival and Growth Test for Sediments

12.1 Introduction

12.1.1 *Chironomus tentans* (Fabricius) have many desirable characteristics of an ideal sediment toxicity testing organism including relative sensitivity to contaminants associated with sediment, contact with sediment, ease of culture in the laboratory, tolerance to varying physicochemical characteristics of sediment, and short generation time. Their response has been evaluated in interlaboratory studies and has been confirmed with natural benthic populations. Many investigators have successfully used *C. tentans* to evaluate the toxicity of freshwater sediments (e.g., Wentsel et al., 1977; Nebeker et al., 1984a; Nebeker et al., 1988; Adams et al., 1985; Giesy et al., 1988; Hoke et al., 1990; West et al., 1993; Ankley et al., 1993; Ankley et al., 1994a; Ankley et al.,1994b). *C. tentans* has been used for a variety of sediment assessments (West et al., 1993; Hoke et al., 1994, 1995; West et al., 1994; Ankley et al., 1994c). Endpoints typically monitored in 10-d sediment toxicity tests with *C. tentans* include survival and growth (ASTM, 1999a).

12.1.2 A specific test method for conducting a 10-d sediment toxicity test is described in Section 12.2 for *C. tentans.* Methods outlined in Appendix A of USEPA (1994a) and in Section 12.1.1 were used for developing test method 100.2. Results of tests using procedures different from the procedures described in Section 12.2 may not be comparable and these different procedures may alter contaminant bioavailability. Comparison of results obtained using modified versions of these procedures might provide useful information concerning new concepts and procedures for conducting sediment tests with aquatic organisms. If tests are conducted with procedures different from the procedures described in this manual, additional tests are required to determine comparability of results (Section 1.3).

12.2 Recommended Test Method for Conducting a 10-d Sediment Toxicity Test with *Chironomus tentans*

12.2.1 Recommended conditions for conducting a 10-d sediment toxicity test with *C. tentans* are summarized in Table 12.1. A general activity schedule is outlined in Table 12.2. Decisions concerning the various aspects of experimental design, such as the number of treatments, number of test chambers/treatment, and water-quality characteristics should be based on the purpose of the test and the methods of data analysis (Section 16). The number of replicates and concentrations tested depends in part on the significance level selected and the type of statistical analysis. When variability remains constant, the sensitivity of a test increases as the number of replicates increases.

12.2.2 The recommended 10-d sediment toxicity test with *C. tentans* must be conducted at 23°C with a 16L:8D photoperiod at an illuminance of about 100 to 1000 lux (Table 12.1). Test chambers are 300-mL high-form lipless beakers containing 100 mL of sediment and 175 mL of overlying water. Ten second- to third-instar midges (about 10-d old) are used to start a test (Section 10.4.1). The number of replicates/treatment depends on the objective of the test. Eight replicates are recommended for routine testing (see Section 16). Midges in each test chamber are fed 1.5 mL of a 4-g/L Tetrafin® suspension daily. Each test chamber receives 2 volume additions/d of overlying water. Water renewals may be manual or automated. Appendix A describes water-renewal systems that can be used to deliver overlying water. Overlying water can be culture water, well water, surface water, site water, or reconstituted water. For site-specific evaluations, the characteristics of the overlying water should be as similar as possible to the site where sediment is collected. Requirements for test acceptability are summarized in Table 12.3.

12.3 General Procedures

12.3.1 Sediment into Test Chambers

The day before the sediment test is started (Day -1) each sediment should be thoroughly homogenized and added to the test chambers (Section 8.3.1). Sediment should be visually inspected to judge the extent of homogeneity. Excess water on the surface of the sediment can indicate separation of solid and liquid components. If a quantitative measure of homogeneity is required, replicate subsamples should be taken from the sediment batch and analyzed for TOC, chemical concentrations, and particle size.

12.3.1.1 Each test chamber should contain the same amount of sediment, determined either by volume or by weight. Overlying water is added to the chambers in a manner that minimizes suspension of sediment. This can be accomplished by gently pouring water along the sides

Table 12.1 Recommended Test Conditions for Conducting a 10-d Sediment Toxicity Test with *Chironomus tentans*

Parameter	Conditions
1. Test type:	Whole-sediment toxicity test with renewal of overlying water
2. Temperature:	23 ± 1°C
3. Light quality:	Wide-spectrum fluorescent lights
4. Illuminance:	About 100 to 1000 lux
5. Photoperiod:	16L:8D
6. Test chamber:	300-mL high-form lipless beaker
7. Sediment volume:	100 mL
8. Overlying water volume:	175 mL
9. Renewal of overlying water:	2 volume additions/d (Appendix A); continuous or intermittent (e.g., one volume addition every 12 h)
10. Age of organisms:	Second- to third-instar larvae (about 10-d-old larvae; all organisms must be third instar or younger with at least 50% of the organisms at third instar; Section 10.4.1)
11. Number of organisms/chamber:	10
12. Number of replicate chambers/treatment:	Depends on the objective of the test. Eight replicates are recommended for routine testing (see Section 16).
13. Feeding:	Tetrafin® goldfish food, fed 1.5 mL daily to each test chamber (1.5 mL contains 6.0 mg of dry solids)
14. Aeration:	None, unless dissolved oxygen in overlying water drops below 2.5 mg/L.
15. Overlying water:	Culture water, well water, surface water, site water, or reconstituted water
16. Test chamber cleaning:	If screens become clogged during a test, gently brush the *outside* of the screen (Appendix A).
17. Overlying water quality:	Hardness, alkalinity, conductivity, pH, and ammonia at the beginning and end of a test. Temperature and dissolved oxygen daily.
18. Test duration:	10 d
19. Endpoints:	Survival and growth (ash-free dry weight, AFDW)
20. Test acceptability:	Minimum mean control survival must be 70%, with minimum mean weight/ surviving control organism of 0.48 mg AFDW. Performance-based criteria specifications are outlined in Table 12.3.

of the chambers or by pouring water onto a baffle (e.g., a circular piece of Teflon with a handle attached) placed above the sediment to dissipate the force of the water. Renewal of overlying water is started on Day -1. A test begins when the organisms are added to the test chambers (Day 0).

12.3.2 Renewal of Overlying Water

12.3.2.1 Renewal of overlying water is required during a test. At any particular time during the test, flow rates through any two test chambers should not differ by more than 10%. Hardness, alkalinity and ammonia concentrations in the water above the sediment, within a treatment, typically should not vary by more than 50% during the test. Mount and Brungs (1967) diluters have been modified for sediment testing, and other automated water-delivery systems have also been used (Maki, 1977; Ingersoll and Nelson, 1990; Benoit et al., 1993; Zumwalt et al., 1994; Brunson et al., 1998; Wall et al., 1998; Leppanen and Maier, 1998). Each water-delivery system should be calibrated before a test is started to verify that the system is functioning properly. Renewal of overlying water is started on Day -1 before the addition of test organisms or food on Day 0. Appendix A describes water-renewal systems that can be used for conducting sediment tests.

12.3.2.2 In water-renewal tests with one to four volume additions of overlying water/d, water-quality characteristics generally remain similar to the inflowing water (Ingersoll and Nelson, 1990; Ankley et al., 1993); however, in static tests, water quality may change profoundly during the exposure (Shuba et al., 1978). For example, in static whole-sediment tests, the alkalinity, hardness, and conductivity of overlying water more than doubled in several treatments during a four-week exposure (Ingersoll and Nelson, 1990). Additionally, concentrations of metabolic products (e.g., ammonia) may also increase during static exposures, and these compounds can either be directly toxic to the test organisms or may contribute to the toxicity of the contaminants in the sediment. Furthermore, changes in water-quality characteristics such as hardness may influence the toxicity of many inorganic (Gauss et

Table 12.2 General Activity Schedule for Conducting a 10-d Sediment Toxicity Test with *Chironomus tentans* [1]

Day	Activity
-14	Isolate adults for production of egg cases.
-13	Place newly deposited egg cases into hatching dishes.
-12	Prepare a larval rearing chamber with new substrate.
-11	Examine egg cases for hatching success. If egg cases have hatched, transfer first-instar larvae and any remaining unhatched embryos from the crystallizing dishes into the larval rearing chamber. Feed organisms.
-10	Same as Day -11.
-9 to -2	Feed and observe midges (Section 10.4). Measure water quality (e.g., temperature and dissolved oxygen).
-1	Add food to each larval rearing chamber and measure temperature and dissolved oxygen. Add sediment into each test chamber, place chamber into exposure system, and start renewing overlying water.
0	Measure total water quality (temperature, pH, hardness, alkalinity, dissolved oxygen, conductivity, ammonia). Remove third-instar larvae from the culture chamber substrate. Add 1.5 mL of Tetrafin® (4.0 g/L) into each test chamber. Transfer 10 larvae into each test chamber. Release organisms under the surface of the water. Archive 20 test organisms for instar determination and weight or length determination. Observe behavior of test organisms.
1 to 8	Add 1.5 mL of food to each test chamber. Measure temperature and dissolved oxygen. Observe behavior of test organisms.
9	Measure total water quality.
10	Measure temperature and dissolved oxygen. End the test by collecting the midges with a sieve. Measure weight or length of surviving larvae.

[1] Modified from Call et al., 1994

al., 1985) and organic (Mayer and Ellersieck, 1986) contaminants. Although contaminant concentrations are reduced in the overlying water in water-renewal tests, organisms in direct contact with sediment generally receive a substantial proportion of a contaminant dose directly from either the whole sediment or from the interstitial water.

12.3.3 Acclimation

12.3.3.1 Test organisms must be cultured and tested at 23°C. Ideally, test organisms should be cultured in the same water that will be used in testing. However, acclimation of test organisms to the test water is not required.

12.3.3.2 Culturing of organisms and toxicity assessment are typically conducted at 23°C. However, occasionally there is a need to perform evaluations at temperatures different than that recommended. Under these circumstances, it may be necessary to acclimate organisms to the desired test temperature to prevent thermal shock when moving immediately from the culture temperature to the test temperature (ASTM, 1999a). Acclimation can be achieved by exposing organisms to a gradual change in temperature; however, the rate of change should be relatively slow to prevent thermal shock. A change in temperature of 1°C every 1 to 2 h has been used successfully in some studies (P.K. Sibley, University of Guelph, Guelph, Ontario, personal communication; APHA, 1989). Testing at temperatures other than 23°C needs to be preceded by studies to determine expected performance under alternate conditions.

12.3.4 Placing Organisms in Test Chambers

12.3.4.1 Test organisms should be handled as little as possible. Midges should be introduced into the overlying water below the air-water interface. Test organisms can be pipetted directly into overlying water. Developmental stage of the test organisms should be documented from a subset of at least 20 organisms used to start the test (Section 10.4.1). Developmental stage can be determined from head capsule width (Table 10.2), length (4 to 6 mm), or dry weight (0.08 to 0.23 mg/individual). It is desirable to measure size at test initiation using the same measure as will be used to assess growth at the end of the test.

12.3.5 Feeding

12.3.5.1 For each beaker, 1.5 mL of Tetrafin® is fed from Day 0 to Day 9. Without addition of food, the test organisms may starve during exposures. However, the addition of the food may alter the availability of the contaminants in the sediment (Wiederholm et al., 1987; Harkey et al., 1994). Furthermore, if too much food is added to the test chamber or if the mortality of test organisms is high, fungal or bacterial growth may develop on the sediment surface. Therefore, the amount of food added to the test chambers is kept to a minimum.

12.3.5.2 Suspensions of food should be thoroughly mixed before aliquots are taken. If excess food collects on the sediment, a fungal or bacterial growth may develop on the sediment surface, in which case feeding should be suspended for one or more days. A drop in dissolved oxygen below 2.5 mg/L during a test may indicate that the food

Table 12.3 Test Acceptability Requirements for a 10-d Sediment Toxicity Test with *Chironomus tentans*

A. It is recommended for conducting a 10-d test with *C. tentans* that the following performance criteria be met:

1. Tests must be started with second- to third-instar larvae (about 10-d-old larvae; see Section 10.4.1).
2. Average survival of *C. tentans* in the control sediment must be greater than or equal to 70% at the end of the test.
3. Average size of *C. tentans* in the control sediment must be at least 0.48 mg AFDW at the end of the test.
4. Hardness, alkalinity, and ammonia in the overlying water typically should not vary by more than 50% during the test, and dissolved oxygen should be maintained above 2.5 mg/L in the overlying water.

B. Performance-based criteria for culturing *C. tentans* include the following:

1. It may be desirable for laboratories to periodically perform 96-h water-only reference-toxicity tests to assess the sensitivity of culture organisms (Section 9.16.2). Data from these reference-toxicity tests could be used to assess genetic strain or life-stage sensitivity of test organisms to select chemicals.
2. Laboratories should keep a record of time to first emergence for each culture and record this information using control charts. Records should also be kept on the frequency of restarting cultures.
3. Laboratories should record the following water-quality characteristics of the cultures at least quarterly: pH, hardness, alkalinity, and ammonia. Dissolved oxygen in the cultures should be measured weekly. Temperature of the cultures should be recorded daily. If static cultures are used, it may be desirable to measure water quality more frequently.
4. Laboratories should characterize and monitor background contamination and nutrient quality of food if problems are observed in culturing or testing organisms.
5. Physiological measurements such as lipid content might provide useful information regarding the health of the cultures.

C. Additional requirements:

1. All organisms in a test must be from the same source.
2. Storage of sediments collected from the field should follow guidance outlined in Section 8.2.
3. All test chambers (and compartments) should be identical and should contain the same amount of sediment and overlying water.
4. Negative-control sediment and appropriate solvent controls must be included in a test. The concentration of solvent used must not adversely affect test organisms.
5. Test organisms must be cultured and tested at 23°C (±1°C).
6. The daily mean test temperature must be within ±1°C of 23°C. The instantaneous temperature must always be within ±3°C of 23°C.
7. Natural physico-chemical characteristics of test sediment collected from the field should be within the tolerance limits of the test organisms.

added is not being consumed. Feeding should be suspended for the amount of time necessary to increase the dissolved oxygen concentration (ASTM, 1999a). If feeding is suspended in one treatment, it should be suspended in all treatments. Detailed records of feeding rates and the appearance of the sediment surface should be made daily.

12.3.6 Monitoring a Test

12.3.6.1 All chambers should be checked daily and observations made to assess test organism behavior such as sediment avoidance. However, monitoring effects on burrowing activity of test organisms may be difficult because the test organisms are often not visible during the exposure. The operation of the exposure system should be monitored daily.

12.3.6.2 Measurement of Overlying Water-Quality Characteristics

12.3.6.2.1 Conductivity, hardness, pH, alkalinity, and ammonia should be measured in all treatments at the beginning and end of a test. Overlying water should be sampled just before water renewal from about 1 to 2 cm above the sediment surface using a pipet. It may be necessary to composite water samples from individual replicates. The pipet should be checked to make sure no organisms are removed during sampling of overlying water. Water quality should be measured on each batch of water prepared for the test.

12.3.6.2.2 Water-only exposures evaluating the tolerance of *C. tentans* larvae to depressed DO have indicated that significant reductions in weight occurred after 10-d exposure to 1.1 mg/L DO, but not at 1.5 mg/L (V. Mattson, USEPA, Duluth, MN, personal communication). This finding concurs with the observations during method development at the USEPA laboratory in Duluth that excursions of DO as low as 1.5 mg/L did not seem to have an effect on midge survival and development (P.K. Sibley, University of Guelph, Guelph, Ontario, personal communication). Based on these findings, it appears that periodic depressions of DO below 2.5 mg/L (but not below 1.5 mg/L) are not likely to adversely affect test results, and thus should not be a reason to discard test data. Nonetheless, tests should be managed toward a goal of DO > 2.5 mg/L to insure satisfactory performance. If the

DO level of the water falls below 2.5 mg/L for any one treatment, aeration is encouraged and should be done in all replicates for the duration of the test. Occasional brushing of screens on outside of beakers will help maintain the exchange of water during renewals using the exposure system described by Benoit et al. (1993). If a probe is used to measure DO in overlying water, it should be thoroughly inspected between samples to make sure that organisms are not attached and should be rinsed between samples to minimize cross contamination. Aeration can be used to maintain dissolved oxygen in the overlying water above 2.5 mg/L (i.e., about 1 bubble/ second in the overlying water).

12.3.6.2.3 Temperature should be measured at least daily in at least one test chamber from each treatment. The temperature of the water bath or the exposure chamber should be continuously monitored. The daily mean test temperature must be within ±1°C of 23°C. The instantaneous temperature must always be within ±3°C of 23°C.

12.3.7 Ending a Test

12.3.7.1 Immobile organisms isolated from the sediment surface or from sieved material should be considered dead. A #40 sieve (425-µm mesh) can be used to remove midges from sediment. Alternatively, Kemble et al. (1994) suggest sieving of sediment using the following procedure: (1) pour about half of the overlying water through a #50- (300-µm) U.S. standard mesh sieve, (2) pour about half of the sediment through the #50-mesh sieve and wash the contents of the sieve into an examination pan, (3) rinse the coarser sediment remaining in the test chamber through a #40- (425-µm) mesh sieve and wash the contents of this second sieve into a second examination pan. Surviving midges can then be isolated from these pans. See Section 12.3.8.1 and 12.3.8.2 for the procedures for measuring weight or length of midges.

12.3.7.2 A consistent amount of time should be taken to examine sieved material for recovery of test organisms (e.g., 5 min/replicate). Laboratories should demonstrate that their personnel are able to recover an average of at least 90% of the organisms from whole sediment. For example, test organisms could be added to control sediment and recovery could be determined after 1 h (Tomasovic et al., 1994).

12.3.8 Test Data

12.3.8.1 Ash-free dry weight (AFDW) and survival are the endpoints measured at the end of the 10-d sediment toxicity test with *C. tentans*. The 10-d method for *C. tentans* in the first edition of this manual (USEPA, 1994a), as well as most previous research, has used dry weight as a measure of growth. However, Sibley et al. (1997b) found that the grain size of sediments influences the amount of sediment that *C. tentans* larvae ingest and retain in their gut. As a result, in finer-grain sediments, a substantial portion of the measured dry weight may be comprised of sediment rather than tissue. While this may not represent a strong bias in tests with identical grain size distributions in all treatments, most field assessments are likely to have varying grain size among sites. This will likely create differences in dry weight among treatments that are not reflective of true somatic growth. For this reason, weight of midges should be measured as ash-free dry weight (AFDW) instead of dry weight. AFDW will more directly reflect actual differences in tissue weight by reducing the influence of sediment in the gut. The duration of the 10-d test starting with third-instar larvae is not long enough to determine emergence of adults. Average size of *C. tentans* in the control sediment must be at least 0.6 mg at the end of the test (0.48 mg AFDW) (Ankley et al., 1993; ASTM, 1999a; Section 17.5). If test organisms are to be used for an evaluation of bioaccumulation, it is not advisable to dry the sample before conducting the residue analysis. If conversion from wet weight to dry weight is necessary, aliquots of organisms can be weighed to establish wet to dry weight conversion factors. A consistent procedure should be used to remove the excess water from the organisms before measuring wet weight.

12.3.8.2 For determination of AFDW, first pool all living larvae in each replicate and dry the sample to a constant weight (e.g., 60°C for 24 h). Note that the weigh boats should be ashed before use to eliminate weighing errors due to the pan oxidizing during ashing. The sample is brought to room temperature in a dessicator and weighed to the nearest 0.01 mg to obtain mean weights per surviving organism per replicate. The dried larvae in the pan are then ashed at 550°C for 2 h. The pan with the ashed larvae is then reweighed and the tissue mass of the larvae is determined as the difference between the weight of the dried larvae plus pan and the weight of the ashed larvae plus pan. In rare instances where preservation is required, an 8% sugar formalin solution can be used to preserve samples (USEPA, 1994a), but the effects of preservation on the weights and lengths of the midges have not been sufficiently studied. Pupae or adult organisms must not be included in the sample to estimate ash-free dry weight. If head capsule width is to be measured, it should be measured on surviving midges at the end of the test before ash-free dry weight is determined.

12.3.8.3 Measurement of length is optional. Separate replicate beakers should be set up to sample lengths of midges at the end of an exposure. An 8% sugar formalin solution can be used to preserve samples for length measurements (Ingersoll and Nelson, 1990). The sugar formalin solution is prepared by adding 120 g of sucrose to 80 mL of formalin, which is then brought to a volume of 1 L using deionized water. This stock solution is mixed with an equal volume of deionized water when used to preserve organisms. NoTox® (Earth Safe Industries, Belle Mead, NJ) can be used as a substitute for formalin (Unger et al., 1993). Midge body length (±0.1 mm) can be measured from the anterior of the labrum to the posterior of the last abdominal segment (Smock, 1980). Kemble et al. (1994) photographed midges at magnification of 3.5X and measured the images using a computer-interfaced digitizing tablet. A digitizing system and microscope can

also be used to measure length (Ingersoll and Nelson, 1990).

12.4 Interpretation of Results

12.4.1 Section 16 describes general information for interpretation of test results. The following sections describe species-specific information that is useful in helping to interpret the results of sediment toxicity tests with *C. tentans*.

12.4.2 Age Sensitivity

12.4.2.1 Midges are perceived to be relatively insensitive organisms in toxicity assessments (Ingersoll, 1995). This conclusion is based on measuring survival of fourth-instar larvae in short-term water-only exposures, a procedure that may underestimate the sensitivity of midges to toxicants. The first and second instars of chironomids are more sensitive to contaminants than the third or fourth instars. For example, first-instar *C. tentans* larvae were 6 to 27 times more sensitive than fourth-instar larvae to acute copper exposure (Nebeker et al., 1984b; Gauss et al., 1985; Figure 12.1) and first-instar *C. riparius* larvae were 127 times more sensitive than second-instar larvae to acute cadmium exposure (Williams et al., 1986b; Figure 12.1). In chronic tests with first-instar larvae, midges were often as sensitive as daphnids to inorganic and organic compounds (Ingersoll et al., 1990). Sediment tests should be started with uniform age and size midges because of the dramatic differences in sensitivity of midges by age. Whereas third-instar midges are not as sensitive as younger organisms, the larger larvae are easier to handle and isolate from sediment at the end of a test.

12.4.2.2 DeFoe and Ankley (1998) studied a variety of contaminated sediments and showed that the sensitivity of *C. tentans* 10-d tests is greatly increased by measurement of growth in addition to survival. Growth of midges in 10-d sediment tests was found to be a more sensitive endpoint than survival of *Hyalella azteca* (DeFoe and Ankley, 1998). In cases where sensitivity of organisms before the third instar is of interest, the long-term sediment exposures can be used, since they begin with newly hatched larvae (Section 15).

12.4.3 Physical characteristics of sediment

12.4.3.1 Grain Size

12.4.3.1.1 Larvae of *C. tentans* appear to be tolerant of a wide range of particle size conditions in substrates. Several studies have shown that survival is not affected by particle size in natural sediments, sand substrates, or formulated sediments in both 10-d and long-term exposures (Ankley et al., 1994; Suedel and Rodgers, 1994; Sibley et al., 1997b, 1998). Ankley et al. (1994a) found that growth of *C. tentans* larvae was weakly correlated with sediment grain size composition, but not organic carbon, in 10-d tests using 50 natural sediments from the Great Lakes. However, Sibley et al. (1997b) found that the correlation between grain size and larval growth disappeared after accounting for inorganic material contained within larval guts and concluded that growth of *C. tentans* was not related to grain size composition in either natural sediments or sand substrates. Avoiding confounding influences of gut contents on weight is the impetus for recommending ash-free dry weight (instead of dry weight) as the index of growth in the 10-day and long-term *C. tentans* tests. Failing to do so could lead to erroneous conclusions regarding the toxicity of the test sediment (Sibley et al., 1997b). Procedures for correcting for gut contents are described in Section 12.3.8. Emergence, reproduction (mean eggs/female), and hatch success were also not affected by the particle size composition of substrates in long-term tests with *C. tentans* (Sibley et al., 1998; Section 15).

12.4.3.2 Organic Matter

12.4.3.2.1 Based on 10-d tests, the content of organic matter in sediments does not appear to affect survival of *C. tentans* larvae in natural and formulated sediments, but may be important with respect to larval growth. Ankley et al. (1994a) found no relationship between sediment organic content and survival or growth in 10-d bioassays with *C. tentans* in natural sediments. Suedel and Rodgers (1994) observed reduced survival in 10-d tests with a formulated sediment when organic matter was <0.91%; however, supplemental food was not supplied in this study, which may influence these results relative to the 10-d test procedures described in this manual. Lacey et al. (1999) found that survival of *C. tentans* larvae was generally not affected in 10-d tests by either the quality or quantity of synthetic (alpha-cellulose) or naturally derived (peat, maple leaves) organic material spiked into a formulated sediment, although a slight reduction in survival below the acceptability criterion (70%) was observed in a natural sediment diluted with formulated sediment at an organic matter content of 6%. In terms of larval growth, Lacey et al. (1999) did not observe any systematic relationship between the level of organic material (e.g., food quantity) and larval growth for each carbon source. Although a significant reduction in growth was observed at the highest concentration (10%) of the leaf treatment in the food quantity study, significantly higher larval growth was observed in this treatment when the different carbon sources were compared at about equal concentrations (effect of food quality). In the latter study, the following gradient of larval growth was established in relation to the source of organic carbon: peat < natural sediment < alpha-cellulose < leaves. Since all of the treatments received a supplemental source of food, these data suggest that both the quality and quantity of organic carbon in natural and formulated sediments may represent an important confounding factor for the growth endpoint in tests with *C. tentans* (Lacey et al., 1999). However, it is important to note that these data are based on 10-d tests; the applicability of these data to long-term testing has not been evaluated (Section 15).

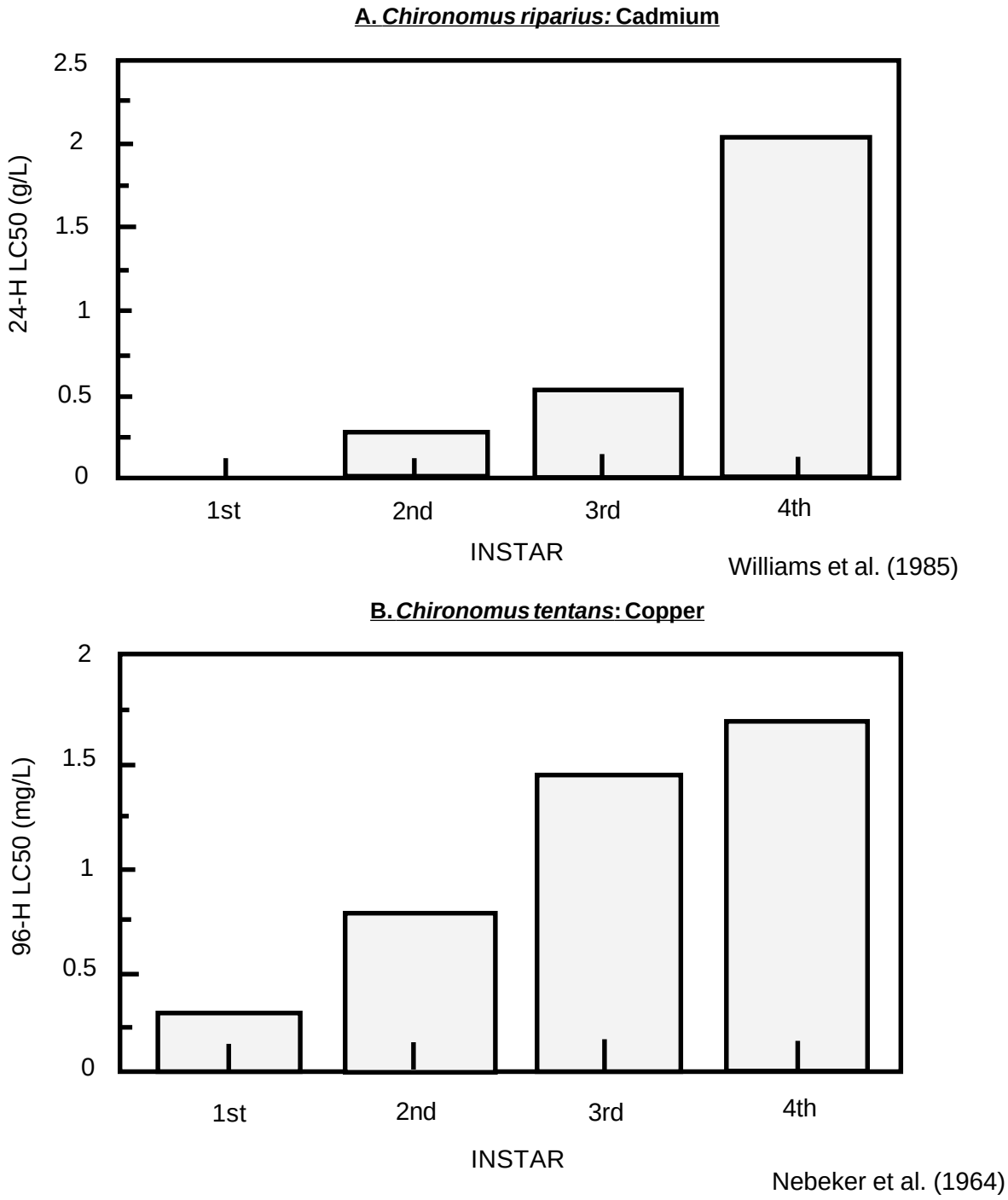


Figure 12.1 Lifestage sensitivity of chironomids.

12.4.4 Isolating Organisms at the End of a Test

12.4.4.1 Quantitative recovery of larvae at the end of a 10-d sediment test should not be a problem. The larvae are red and typically greater than 5 mm long.

12.4.5 Influence of Indigenous Organisms

12.4.5.1 The influence of indigenous organisms on the response of *C. tentans* in sediment tests has not been reported. Survival of a closely related species, *C. riparius* was not reduced in the presence of oligochaetes in sediment samples (Reynoldson et al., 1994). However, growth of *C. riparius* was reduced when high numbers of oligochaetes were placed in a sample. Therefore, it is important to determine the number and biomass of indigenous organisms in field-collected sediment in order to better interpret growth data (Reynoldson et al., 1994; DeFoe and Ankley, 1998). Furthermore, presence of predators may also influence the response of test organisms in sediment (Ingersoll and Nelson, 1990).

12.4.6. Sexual Dimorphism

12.4.6.1 Differences in size between males and females of a closely related midge species (*Chironomus riparius*) had little effect on interpretation of growth-related effects in sediment tests (<3% probability of making a Type I error [nontoxic sample classified as toxic] due to sexual dimorphism; Day et al., 1994). Therefore, sexual dimorphism will probably not be a confounding factor when interpreting growth results measured in sediment tests with *C. tentans*.

12.4.7 Ammonia Toxicity

12.4.7.1 Section 1.3.7.5 addresses interpretative guidance for evaluating toxicity associated with ammonia in sediment.

Section 13
Test Method 100.3
Lumbriculus variegatus Bioaccumulation Test for Sediments

13.1 Introduction

13.1.1 *Lumbriculus variegatus* (Oligochaeta) have many desirable characteristics of an ideal sediment bioaccumulation testing organism including contact with sediment, ease of culture in the laboratory, and tolerance to varying physico-chemical characteristics of sediment. The response of *L. variegatus* in laboratory exposures has been confirmed with natural benthic populations. Many investigators have successfully used *L. variegatus* in toxicity or bioaccumulation tests. Toxicity studies have been conducted in water-only tests (Bailey and Liu, 1980; Hornig, 1980; Ewell et al., 1986; Nebeker et al., 1989; Ankley et al., 1991a; Ankley et al., 1991b), in effluent tests (Hornig, 1980), and in whole-sediment tests (Nebeker et al., 1989; Ankley et al., 1991a; Ankley et al., 1991b; Ankley et al., 1992a; Call et al., 1991; Carlson et al., 1991; Phipps et al., 1993; West et al., 1993). Several studies have reported the use of *L. variegatus* to examine bioaccumulation of chemicals from sediment (Schuytema et al., 1988; Nebeker et al., 1989; Ankley et al., 1991b; Call et al., 1991; Carlson et al., 1991; Ankley et al., 1993; Kukkonen and Landrum, 1994; and Brunson et al., 1993, 1998). However, interlaboratory studies have not yet been conducted with *L. variegatus.*

13.1.2 Additional research is needed on the standardization of bioaccumulation procedures with sediment. Therefore, Section 13.2 describes general guidance for conducting a 28-d sediment bioaccumulation test with *L. variegatus*. Methods outlined in Appendix A of USEPA (1994a) and in Section 13.1.1 were used for developing this general guidance. Results of tests using procedures different from the procedures described in Section 13.2 may not be comparable, and these different procedures may alter bioavailability. Comparison of results obtained using modified versions of these procedures might provide useful information concerning new concepts and procedures for conducting sediment tests with aquatic organisms. If tests are conducted with procedures different from the procedures described in this manual, additional tests are required to determine comparability of results (Section 1.3).

13.2 Procedure for Conducting Sediment Bioaccumulation Tests with *Lumbriculus variegatus*

13.2.1 Recommended test conditions for conducting a 28-d sediment bioaccumulation test with *L. variegatus* are summarized in Table 13.1. Table 13.2 outlines procedures for conducting sediment toxicity tests with *L. variegatus*. A general activity schedule is outlined in Table 13.3. Decisions concerning the various aspects of experimental design, such as the number of treatments, number of test chambers/treatment, and water-quality characteristics should be based on the purpose of the test and the methods of data analysis (Section 16). The number of replicates and concentrations tested depends in part on the significance level selected and the type of statistical analysis. When variability remains constant, the sensitivity of a test increases as the number of replicates increases.

13.2.2 The recommended 28-d sediment bioaccumulation test with *L. variegatus* can be conducted with adult oligochaetes at 23°C with a 16L:8D photoperiod at a illuminance of about 100 to 1000 lux (Table 13.1). Test chambers can be 4 to 6 L that contain 1 to 2 L of sediment and 1 to 4 L of overlying water. The number of replicates/treatment depends on the objective of the test. Five replicates are recommended for routine testing (Section 16). To minimize depletion of sediment contaminants, the ratio of total organic carbon in sediment to dry weight of organisms should be about 50:1. A minimum of 1 g/replicate with up to 5 g/replicate should be tested. Oligochaetes are not fed during the test. Each chamber receives 2 volume additions/d of overlying water. Appendix A and Brunson et al., (1998) describe water-renewal systems that can be used to deliver overlying water. Overlying water can be culture water, well water, surface water, site water, or reconstituted water. For site-specific evaluations, the characteristics of the overlying water should be as similar as possible to the site where sediment is collected. Requirements for test acceptability are outlined in Table 13.4.

13.2.2.1 Before starting a 28-d sediment bioaccumulation test with *L. variegatus,* a toxicity screening test can be conducted for at least 4 d using procedures outlined in Table 13.2 (Brunson et al., 1993). The preliminary toxicity screening test is conducted at 23°C with a 16L:8D photoperiod at an illuminance of about 100 to 1000 lux. Test chambers are 300-mL high-form lipless beakers containing

Table 13.1 Recommended Test Conditions for Conducting a 28-d Sediment Bioaccumulation Test with *Lumbriculus variegatus*

Parameter	Conditions
1. Test type:	Whole-sediment bioaccumulation test with renewal of overlying water
2. Temperature:	23 ± 1°C
3. Light quality:	Wide-spectrum fluorescent lights
4. Illuminance:	About 100 to 1000 lux
5. Photoperiod:	16L:8D
6. Test chamber:	4- to 6-L aquaria with stainless steel screens or glass standpipes
7. Sediment volume:	1 L or more depending on TOC
8. Overlying water volume:	1 L or more depending on TOC
9. Renewal of overlying water:	2 volume additions/d (Appendix A); continuous or intermittent (e.g., one volume addition every 12 h)
10. Age of test organisms:	Adults
11. Loading of organisms in chamber:	Ratio of total organic carbon in sediment to organism dry weight should be no less than 50:1. Minimum of 1 g/replicate. Preferably 5 g/replicate.
12. Number of replicate chambers/treatment:	Depends on the objective of the test. Five replicates are recommended for routine testing (see Section 16).
13. Feeding:	None
14. Aeration:	None, unless DO in overlying water drops below 2.5 mg/L
15. Overlying water:	Culture water, well water, surface water, site water, or reconstituted water
16. Test chamber cleaning:	If screens become clogged during the test, gently brush the *outside* of the screen (Appendix A).
17. Overlying water quality:	Hardness, alkalinity, conductivity, pH, and ammonia at the beginning and end of a test. Temperature and dissolved oxygen daily.
18. Test duration:	28 d
19. Endpoint:	Bioaccumulation
20. Test acceptability:	Performance-based criteria specifications are outlined in Table 13.4.

100 mL of sediment and 175 mL of overlying water. Ten adult oligochaetes/replicate are used to start a test. Four replicates are recommended for toxicity screening tests. Oligochaetes are not fed during the test. Each chamber receives 2 volume additions/d of overlying water. Appendix A and Brunson et al. (1998) describe water-renewal systems that can be used to deliver overlying water. Overlying water should be similar to the water to be used in the bioaccumulation test. Endpoints monitored at the end of a toxicity test are number of organisms and behavior. Numbers of *L. variegatus* in the toxicity screening test should not be significantly reduced in the test sediment relative to the control sediment. Test organisms should burrow into test sediment. Avoidance of test sediment by *L. variegatus* may decrease bioaccumulation.

13.3 General Procedures

13.3.1 Sediment into Test Chambers

13.3.1.1 The day before the sediment test is started (Day -1) each sediment should be thoroughly homogenized and added to the test chambers (Section 8.3.1). Sediment should be visually inspected to judge the extent of homogeneity. Excess water on the surface of the sediment can indicate separation of solid and liquid components. If a quantitative measure of homogeneity is required, replicate subsamples should be taken from the sediment batch and analyzed for TOC, chemical concentrations, and particle size.

13.3.1.2 Each test chamber should contain the same amount of sediment, determined either by volume or by weight. Overlying water is added to the chambers in a manner that minimizes suspension of sediment. This can be accomplished by gently pouring water along the sides of the chambers or by pouring water onto a baffle (e.g., a circular piece of Teflon® with a handle attached) placed above the sediment to dissipate the force of the water. Renewal of overlying water is started on Day -1. A test begins when the organisms are added to the test chambers (Day 0).

13.3.2 Renewal of Overlying Water

13.3.2.1 Renewal of overlying water is recommended during a test. At any particular time during the test, flow rates through any two test chambers should not differ by more than 10%. Hardness, alkalinity and ammonia

Table 13.2 Recommended Test Conditions for Conducting a Preliminary 4-d Sediment Toxicity Screening Test with *Lumbriculus variegatus*

	Parameter	Conditions
1.	Test type:	4-d whole-sediment toxicity test with renewal of overlying water
2.	Temperature:	23 ± 1°C
3.	Light quality:	Wide-spectrum fluorescent lights
4.	Illuminance:	About 100 to 1000 lux
5.	Photoperiod:	16L:8D
6.	Test chamber:	300-mL high-form lipless beaker
7.	Sediment volume:	100 mL
8.	Overlying water volume:	175 mL
9.	Renewal of overlying water:	2 volume additions/d (Appendix A); continuous or intermittent (e.g., one volume addition every 12 h)
10.	Age of test organisms:	Adults
11.	Number of organisms/chamber:	10
12.	Number of replicate chambers/treatment:	4 minimum
13.	Feeding:	None
14.	Aeration:	None, unless DO in overlying water drops below 2.5 mg/L
15.	Overlying water:	Culture water, well water, surface water, site water, or reconstituted water
16.	Test chamber cleaning:	If screens become clogged during the test, gently brush the *outside* of the screen.
17.	Overlying water quality:	Hardness, alkalinity, conductivity, pH, and ammonia at the beginning and end of a test. Temperature and dissolved oxygen daily.
18.	Test duration:	4 d (minimum; up to 10 d)
19.	Endpoints:	Number of organisms and behavior. There should be no significant reduction in number of organisms in a test sediment relative to the control.
20.	Test acceptability:	Performance-based criteria specifications are outlined in Table 13.4.

concentrations in the water above the sediment, within a treatment, should not vary by more than 50% during the test. Mount and Brungs (1967) diluters have been modified for sediment testing, and other automated water-delivery systems have also been used (Maki, 1977; Ingersoll and Nelson, 1990; Benoit et al., 1993; Zumwalt et al., 1994; Brunson et al., 1998; Wall et al., 1998; Leppanen and Maier, 1998). Each water-delivery system should be calibrated before a test is started to verify that the system is functioning properly. Renewal of overlying water is started on Day -1 before the addition of test organisms on Day 0 (Appendix A).

13.3.2.2 In water-renewal tests with one to four volume additions of overlying water/d, water-quality characteristics generally remain similar to the inflowing water (Ingersoll and Nelson, 1990; Ankley et al., 1993); however, in static tests, water quality may change profoundly during the exposure (Shuba et al., 1978). For example, in static whole-sediment tests, the alkalinity, hardness, and conductivity of overlying water more than doubled in several treatments during a four-week exposure (Ingersoll and Nelson, 1990). Additionally, concentrations of metabolic products (e.g., ammonia) may also increase during static exposures, and these compounds can either be directly toxic to the test organisms or may contribute to the toxicity of the contaminants in the sediment. Furthermore, changes in water-quality characteristics such as hardness may influence the toxicity of many inorganic (Gauss et al., 1985) and organic (Mayer and Ellersieck, 1986) contaminants. Although contaminant concentrations are reduced in the overlying water in water-renewal tests, organisms in direct contact with sediment generally receive a substantial proportion of a contaminant dose directly from either the whole sediment or from the interstitial water.

13.3.3 Acclimation

13.3.3.1 Test organisms must be cultured and tested at 23°C. Ideally, test organisms should be cultured in the same water that will be used in testing. However, acclimation of test organisms to the test water is not required.

13.3.3.2 Culturing of organisms and toxicity assessment are typically conducted at 23°C. However, occasionally there is a need to perform evaluations at temperatures different than that recommended. Under these circumstances, it may be necessary to acclimate organisms to the desired test temperature to prevent thermal shock

Table 13.3 General Activity Schedule for Conducting a 28-d Sediment Bioaccumulation Test with *Lumbriculus variegatus*

A. Conducting a 4-d Toxicity Screening Test (conducted before the 28-d bioaccumulation test)

Day	*Activity*
-1	Isolate worms for conducting toxicity screening test. Add sediment into each test chamber, place chambers into exposure system, and start renewing overlying water.
0	Measure total water quality (pH, temperature, dissolved oxygen, hardness, alkalinity, conductivity, ammonia). Transfer 10 worms into each test chamber. Measure weight of a subset of 20 organisms used to start the test. Observe behavior of test organisms.
1 to 2	Measure temperature and dissolved oxygen. Observe behavior of test organisms.
3	Same as Day 1. Measure total water quality.
4	Measure temperature and dissolved oxygen. End the test by collecting the oligochaetes with a sieve and determine weight of survivors. Bioaccumulation tests should not be conducted with *L. variegatus* if a test sediment significantly reduces number of oligochaetes relative to the control sediment or if oligochaetes avoid the sediment.

B. Conducting a 28-d Bioaccumulation Test

Day	*Activity*
-1	Isolate worms for conducting bioaccumulation test. Add sediment into each test chamber, place chambers into exposure system, and start renewing overlying water.
0	Measure total water quality (pH, temperature, dissolved oxygen, hardness, alkalinity, conductivity, ammonia). Transfer appropriate amount of worms (based on weight) into each test chamber. Sample a subset of worms used to start the test for residue analyses. Observe behavior of test organisms.
1 to 6	Measure temperature and dissolved oxygen. Observe behavior of test organisms.
7	Same as Day 1. Measure total water quality.
8 to 13	Same as Day 1
14	Same as Day 7
15 to 20	Same as Day 1
21	Same as Day 7
22 to 26	Same as Day 1
27	Measure total water quality.
28	Measure temperature and dissolved oxygen. End the uptake by collecting the worms with a sieve. Separate any indigenous organisms from *L. variegatus.* Determine the weight of survivors. Eliminate the gut contents of surviving worms in water for 6 to 8 h. Longer purging periods (not to exceed 24 hours) may be used if all target analytes have Log K_{ow} >5 (Section 13.3.7.3).

when moving immediately from the culture temperature to the test temperature (ASTM, 1999a). Acclimation can be achieved by exposing organisms to a gradual change in temperature; however, the rate of change should be relatively slow to prevent thermal shock. A change in temperature of 1°C every 1 to 2 h has been used successfully in some studies (P.K. Sibley, University of Guelph, Guelph, Ontario, personal communication). Testing at temperatures other than 23°C needs to be preceded by studies to determine expected performance under alternate conditions.

13.3.4 Placing Organisms in Test Chambers

13.3.4.1 Isolate oligochaetes for starting a test as described in Section 10.5.6. A subset of *L. variegatus* at the start of the test should be sampled to determine starting concentrations of chemicals of concern. Mean group weights should be measured on a subset of at least 100 organisms used to start the test. The ratio of total organic carbon in sediment to dry weight of organisms at the start of the test should be no less than 50:1.

13.3.4.2 Oligochaetes added to each replicate should not be blotted to remove excess water (Section 10.5.6). Oligochaetes can be added to each replicate at about 1.33 X of the target stocking weight (Brunson et al., 1998). This additional 33% should account for the excess weight from water in the sample of nonblotted oligochaetes at the start of the test.

13.3.5 Feeding

13.3.5.1 *Lumbriculus variegatus* should not be fed during a bioaccumulation test.

Table 13.4 Test Acceptability Requirements for a 28-d Sediment Bioaccumulation Test with *Lumbriculus variegatus*

A. It is recommended for conducting a 28-d test with *L. variegatus* that the following performance criteria be met:

1. Numbers of *L. variegatus* in a 4-d toxicity screening test should not be significantly reduced in the test sediment relative to the control sediment.
2. Test organisms should burrow into test sediment. Avoidance of test sediment by *L. variegatus* may decrease bioaccumulation.
3. Hardness, alkalinity, and ammonia in the overlying water typically should not vary by more than 50% during the test, and dissolved oxygen should be maintained above 2.5 mg/L in the overlying water.

B. Performance-based criteria for culturing *L. variegatus* include the following:

1. It may be desirable for laboratories to periodically perform 96-h water-only reference toxicity tests to assess the sensitivity of culture organisms (Section 9.16.2). Data from these reference-toxicity tests could be used to assess genetic strain or life-stage sensitivity of test organisms to select chemicals.
2. Laboratories should monitor the frequency with which the population is doubling in the culture (number of organisms) and record this information using control charts (doubling rate would need to be estimated on a subset of animals from a mass culture). Records should also be kept on the frequency of restarting cultures. If static cultures are used, it may be desirable to measure water quality more frequently.
3. Food used to culture organisms should be analyzed before the start of a test for compounds to be evaluated in the bioaccumulation test.
4. Laboratories should record the following water-quality characteristics of the cultures at least quarterly and the day before the start of a sediment test: pH, hardness, alkalinity, and ammonia. Dissolved oxygen in the cultures should be measured weekly. Temperature of the cultures should be recorded daily.
5. Laboratories should characterize and monitor background contamination and nutrient quality of food if problems are observed in culturing or testing organisms.
6. Physiological measurements such as lipid content might provide useful information regarding the health of the cultures.

C. Additional requirements:

1. All organisms in a test must be from the same source.
2. Storage of sediments collected from the field should follow guidance outlined in Section 8.2.
3. All test chambers (and compartments) should be identical and should contain the same amount of sediment and overlying water.
4. Negative-control sediment and/or the appropriate solvent controls must be included in a test. The concentration of solvent used must not affect test organisms adversely.
5. Test organisms must be cultured and tested at 23°C (±1°C).
6. The daily mean test temperature must be within ±1°C of 23°C. The instantaneous temperature must always be within ±3°C of 23°C
7. Natural physico-chemical characteristics of test sediment collected from the field should be within the tolerance limits of the test organisms.

13.3.6 Monitoring a Test

13.3.6.1 All chambers should be checked daily and observations made to assess test organism behavior such as sediment avoidance. However, monitoring effects on burrowing activity of test organisms may be difficult because the test organisms are often not visible during the exposure. The operation of the exposure system should be monitored daily.

13.3.6.2 Measurement of Overlying Water-quality Characteristics

13.3.6.2.1 Conductivity, hardness, pH, alkalinity, and ammonia should be measured in all treatments at the beginning and end of a test. Overlying water should be sampled just before water renewal from about 1 to 2 cm above the sediment surface using a pipet. It may be necessary to composite water samples from individual replicates. The pipet should be checked to make sure no organisms are removed during sampling of overlying water. Water quality should be measured on each batch of water prepared for the test.

13.3.6.2.2 Dissolved oxygen should be measured daily and should be above 2.5 mg/L. If a probe is used to measure dissolved oxygen in overlying water, it should be thoroughly inspected between samples to make sure that organisms are not attached and should be rinsed between samples to minimize cross contamination. Aeration can be used to maintain dissolved oxygen in the overlying water above 2.5 mg/L (i.e., about 1 bubble/second in the overlying water). Dissolved oxygen and pH can be measured directly in the overlying water with a probe.

13.3.6.2.3 Temperature should be measured at least daily in at least one test chamber from each treatment. The temperature of the water bath or the exposure chamber should be continuously monitored. The daily mean test temperature must be within ±1°C of 23°C. The instantaneous temperature must always be within ±3°C of 23°C.

13.3.7 Ending a Test

13.3.7.1 Sediment at the end of the test can be sieved through a fine-meshed screen sufficiently small to retain the oligochaetes (e.g., U.S. standard sieve #40 (425-µm mesh) or #60 (250-µm mesh). The sieved material should be quickly transferred to a shallow pan to keep oligochaetes from moving through the screen. Immobile organisms should be considered dead.

13.3.7.2 The sediment contribution to the body weight of *Lumbriculus variegatus* is reported to be about 20% of the wet weight and the contribution to chemical concentrations ranges from 0 to 11% in two laboratory studies (Kukkonen and Landrum, 1994; 1995). Analyses by Mount et al. (1998) suggest that under certain conditions substantially larger errors may occur if gut contents are included in samples for tissue analysis. Accordingly, after separating the organisms from the sediment, test animals are held in clean water to allow the worms to purge their guts of sediment. To initiate gut purging, live oligochaetes are transferred from the sieved material to a 1-L beaker containing overlying water only. Oligochaetes should not be placed in clean sediment to eliminate gut contents. Clean sediment can add to the dry weight of the oligochaetes, which would result in a dilution of chemical concentrations on a dry weight basis. Further, purging in clean sediment is thought to accelerate depuration of chemical from tissues (Kukkonen and Landrum, 1994). The elimination beakers may need to be aerated to maintain dissolved oxygen above 2.5 mg/L.

13.3.7.3 The first edition of this manual (USEPA, 1994a) specified a 24-h holding period for gut purging, based on the findings of Call et al. (1991) who reported that *L. variegatus* clear more than 90% of their gut contents in 24 h. Kukkonen and Landrum (1995) reported *L. variegatus* will purge out the intestinal contents in 10 h in water, and more recently, Mount et al. (1999) found that gut purging of *L. variegatus* was essentially complete in only 6 h . Shorter purging periods may be preferable to reduce depuration of chemical from tissue during holding in clean water, particularly for compounds with log K_{ow} < 5 (Figure 13.1). Mount et al. (1999) estimated that after a 6-h purging period, compounds with log K_{ow} > 3.85 would remain at >90% of their initial concentrations, but after 24 h, only compounds with log K_{ow} > 5 would be at >90% of the initial concentration in tissue. For this reason, it is recommended that the purging period last 6 to 8 h. Longer purging periods (not to exceed 24 hours) may be used if all target analytes have log K_{ow} > 5.

13.3.7.4 Field-collected sediments may include indigenous oligochaetes. The behavior and appearance of indigenous oligochaetes are usually different from *L. variegatus.* It may be desirable to test extra chambers without the addition of *L. variegatus* to check for the presence of indigenous oligochaetes in field-collected sediment (Phipps et al., 1993). Bioaccumulation of chemicals by indigenous oligochaetes exposed in the same chamber with introduced *L. variegatus* in a 28-d test has been evaluated (Brunson et al., 1993). Peak concentrations of select PAHs and DDT in this study were similar in the indigenous oligochaetes and *L. variegatus* exposed in the same chamber for 28 d.

13.3.7.5 Care should be taken to isolate at least the minimum amount of tissue mass from each replicate chamber needed for analytical chemistry.

13.3.8 Test Data

13.3.8.1 Sensitivity of tissue analyses is dependent largely on the mass of tissue available and the sensitivity of the analytical procedure. To obtain meaningful results from bioaccumulation tests, it is essential that desired detection limits be established before testing, and that the test design allow for sufficient tissue mass. Tissue masses required for various analyses at selected lower limits of detection are listed in Table 13.5. Detection limits for individual PAHs in tissue are listed in Table 13.6. For most chemicals, a minimum mass of 1 g/replicate (wet weight) and preferably 5 g/replicate (wet weight) should be tested. Again, however, to insure results will be meaningful, required masses for analytes of interest to the study should be specifically evaluated before the study is designed.

13.3.8.2 If an estimate of dry weight is needed, a subsample should be dried to a constant weight at about 60 to 90°C. The sample is brought to room temperature in a desiccator and weighed to the nearest 0.01 mg. *Lumbriculus variegatus* typically contain about 1% lipid (wet weight). It may be desirable to determine ash-free dry weight (AFDW) of oligochaetes instead of dry weight. Measurement of AFDW is recommended over dry weight for *C. tentans* due to the contribution of sediment in the gut to the weight of midge (Section 12.3.8; Sibley et al., 1997b). Additional data are needed to determine the contribution of sediment in the gut of *L. variegatus* to body weight before a definitive recommendation can be made to measure AFDW of oligochaetes routinely.

13.3.8.3 Depending on specific study objectives, total lipids can be measured on a subsample of the total tissue mass of each thawed replicate sample. Gardner et al. (1985) describe procedures for measuring lipids in 1 mg of tissue. Different methods of lipid analysis can yield different results (Randall et al., 1991). The analytical method used for lipid analysis should be calibrated against the chloroform-methanol extraction method described by Folch et al. (1957) and Bligh and Dyer (1959).

13.3.8.3.1 A number of studies have demonstrated that lipids are the major storage site for organic chemicals in a variety of organisms (Roberts et al., 1977; Oliver and Niimi, 1983; de Boer, 1988). Because of the importance of

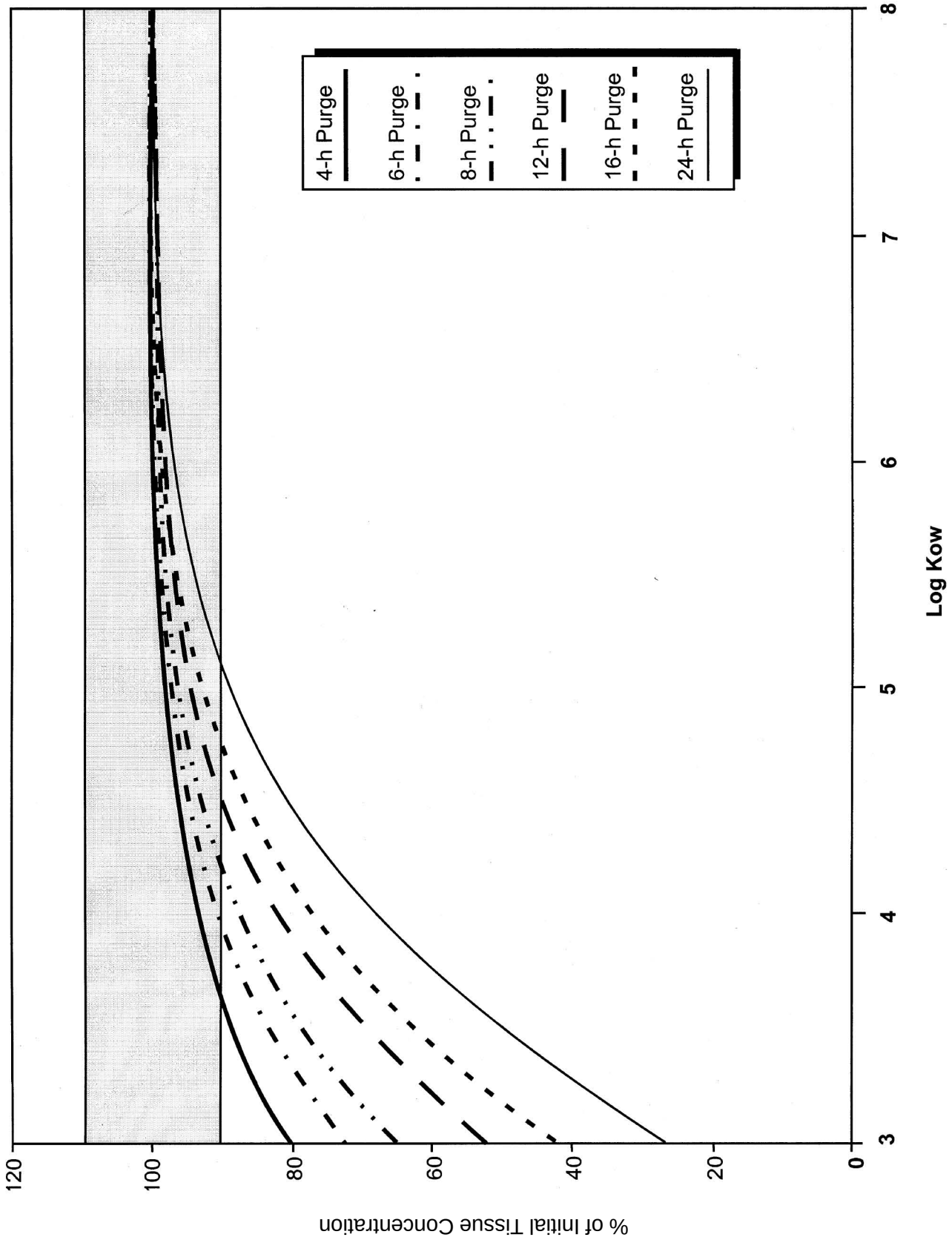


Figure 13.1 Predicted depuration of nonionic organic chemicals from tissue of *Lumbriculus variegatus* as a function of K_{ow} and duration of depuration, assuming no contribution of sediment in the gut. Shaded area represents ±10% of tissue concentration at the beginning of the depuration period (Mount et al., 1999).

Table 13.5 Grams of *Lumbriculus variegatus* Tissue (Wet Weight) Required for Various Analytes at Selected Lower Limits of Detection

Analyte	Grams of Tissue 1.0	2.0	5.0
	Lower Limit of Detection (µg/g)		
PCBs			
PCB (total[1])	0.600	0.300	0.120
PCB (congener[2]) *Level of chlorination*			
mono-trichloro	0.025	0.0125	0.005
tetra-hexachloro	0.050	0.025	0.010
hepta-octachloro	0.075	0.0375	0.015
nona-decachloro	0.125	0.0625	0.025
Organochlorine pesticides [1]			
p,p´ DDE	0.050	0.025	0.010
p,p´ - DDD	0.050	0.025	0.010
p,p´ - DDT	0.050	0.025	0.010
o,p´ - DDE	0.050	0.025	0.010
o,p´ DDD	0.050	0.025	0.010
o,p´ DDT	0.050	0.025	0.010
Alpha-chlordane	0.050	0.025	0.010
Gamma-chlordane	0.050	0.025	0.010
Dieldrin	0.050	0.025	0.010
Endrin	0.050	0.025	0.010
Heptachlorepoxide	0.050	0.025	0.010
Oxychlordane	0.050	0.025	0.010
Mirex	0.050	0.025	0.010
Trans - nonachlor	0.050	0.025	0.010
Toxaphene	0.600	0.300	0.120
PAHs [3]			
PAHs	0.012	0.006	0.002
Dioxins [4]			
TCDD (ng/g)	0.020	0.010	0.004
Inorganic [5]			
Cadmium	0.005	0.0025	0.001
Copper	0.005	0.0025	0.001
Lead	0.005	0.0025	0.001
Zinc	0.005	0.0025	0.001

[1] Schmitt et al., 1990
[2] USEPA, 1990c
[3] Vassilaros et al., 1982
[4] USEPA, 1990d
[5] Schmitt and Finger, 1987

Table 13.6 Detection Limits (ng) of Individual PAHs by HPLC-FD[1]

Analyte	Detection Limit (ng)
Benzo(a)pyrene	0.01
Pyrene	0.03
Benzo(k)fluoranthene	0.03
Dibenz(a,h)anthracene	0.03
Anthracene	0.10
Benz(a)anthracene	0.10
Benzo(e)pyrene	0.10
Benzo(b)fluoranthene	0.10
Benzo(g,h,i)perylene	0.10
3-Methyleholanthrene	0.10

[1] Obana et al., 1981

lipids, it may be desirable to normalize bioaccumulated concentrations of nonpolar organics to the tissue lipid concentration. Lipid concentration is one of the factors required in deriving the BSAF (Section 16). However, the difficulty with using this approach is that each lipid method generates different lipid concentrations (see Kates (1986) for discussion of lipid methodology). The differences in lipid concentrations directly translate to a similar variation in the lipid-normalized chemical concentrations or BSAF.

13.3.8.3.2 For comparison of lipid-normalized tissue residues or BASFs, it is necessary to either promulgate a standard lipid technique or to intercalibrate the various techniques. Standardization of a single method is difficult because the lipid methodology is often intimately tied in with the extraction procedure for chemical analysis. As an interim solution, the Bligh-Dyer lipid method (Bligh and Dyer, 1959) is recommended as a temporary “intercalibration standard” (ASTM, 1999c).

13.3.8.3.3 The potential advantages of Bligh-Dyer include its ability to extract neutral lipids not extracted by many other solvent systems and the wide use of this method (or the same solvent system) in biological and toxicological studies (e.g., Roberts et al., 1977; Oliver and Niimi, 1983; de Boer, 1988; Landrum, 1989). Because the technique is independent of any particular analytical extraction procedure, it will not change when the extraction technique is changed. Additionally, the method can be modified for small tissue sample sizes as long as the solvent ratios are maintained (Herbes and Allen, 1983; Gardner et al., 1985).

13.3.8.3.4 If the Bligh-Dyer method is not the primary lipid method used, the chosen lipid analysis method should be compared with Bligh-Dyer for each tissue type. The chosen lipid method can then be converted to “Bligh-Dyer” equivalents and the lipid-normalized tissue

residues reported in "Bligh-Dyer equivalents." In the interim, it is suggested that extra tissue of each species be frozen for future lipid analysis in the event that a different technique proves more advantageous (ASTM, 1999c).

13.4 Interpretation of Results

13.4.1 Section 16 describes general information for interpretation of test results. The following sections describe species-specific information that is useful in helping to interpret the results of sediment bioaccumulation tests with *L. variegatus.*

13.4.2 Duration of Exposure

13.4.2.1 Because data from bioaccumulation tests often will be used in ecological or human health risk assessments, the procedures are designed to generate quantitative estimates of steady-state tissue residues. Eighty percent of steady state is used as the general criterion (ASTM, 1999c). Because results from a single or few species often will be extrapolated to other species, the procedures are designed to maximize exposure to sediment-associated chemicals so as not to systematically underestimate residues in untested species.

13.4.2.2 A kinetic study can be conducted to estimate steady-state concentrations instead of conducting a 28-d bioaccumulation test (e.g., sample on Day 1, 3, 7, 14, 28; Brunson et al., 1993; USEPA-USACE, 1991). A kinetic test conducted under the same test conditions outlined above, can be used when 80% of steady state will not be obtained within 28 d or when more precise estimates of steady-state tissue residues are required. Exposures shorter than 28 d may be used to determine whether compounds are bioavailable (i.e., bioaccumulation potential).

13.4.2.3 DDT reportedly reached 90% of steady state by Day 14 of a 56-d exposure with *L. variegatus*. However, low molecular weight PAHs (e.g., acenaphthylene, fluorene, phenanthrene) generally peaked at Day 3 and tended to decline to Day 56 (Brunson et al., 1993). In general, concentrations of high molecular weight PAHs (e.g., benzo[b]fluoranthene, benzo[e]pyrene, indeno-[1,2,3-c,d]pyrene) either peaked at Day 28 or continued to increase during the 56-d exposure.

13.4.3 Influence of Indigenous Organisms

13.4.3.1 Field-collected sediments may include indigenous oligochaetes. Phipps et al. (1993) recommend testing extra chambers without the addition of *L. variegatus* to check for the presence of indigenous oligochaetes in field-collected sediment.

13.4.4 Sediment Toxicity in Bioaccumulation Tests

13.4.4.1 Toxicity or altered behavior of organisms in a sample may not preclude use of bioaccumulation data; however, information on adverse effects of a sample should be included in the report.

13.4.4.2 Grain Size.

13.4.4.2.1 *Lumbriculus variegatus* are tolerant of a wide range of substrates. Physico-chemical characteristics (e.g., grain size) of sediment were not significantly correlated to the growth or reproduction of *L. variegatus* in 10-d toxicity tests (see Section 10.1.3.3; Ankley et al., 1994a).

13.4.4.3 Sediment Organic Carbon

13.4.4.3.1 Reduced growth of *L. variegatus* may result from exposure to sediments with low organic carbon concentrations (G.T. Ankley, USEPA, Duluth, MN, personal communication). For this reason, reduced growth observed in bioaccumulation tests could be caused by either direct toxicity or insufficient nutrition of the sediment. Testing additional replicate chambers with supplemental food could be used to help make this distinction, although the effect of added food on accumulation of chemicals would need to be considered in the test interpretation.

13.4.4.4 Ammonia Toxicity

13.4.4.4.1 Section 1.3.7.5 addresses interpretative guidance for evaluating toxicity associated with ammonia in sediment.

Section 14
Test Method 100.4
Hyalella azteca 42-d Test for Measuring the Effects of Sediment-associated Contaminants on Survival, Growth, and Reproduction

14.1 Introduction

14.1.1 *Hyalella azteca* are routinely used to assess the toxicity of chemicals in sediment (Section 11; Nebeker et al., 1984; Dillon and Gibson,1986; Burton et al., 1989; Burton et al., 1992; Ingersoll and Nelson, 1990; Borgmann and Munawar, 1989; Ankley et al., 1994; Winger and Lazier, 1994; Suedel and Rodgers, 1994; Day et al., 1995; Kubitz et al.,1996). Test duration and endpoints recommended in previously developed standard methods for sediment testing with *H. azteca* include 10-d survival (Section 11; USEPA, 1994a) and 10- to 28-d survival and growth (ASTM, 1999a; Environment Canada, 1998a). Short-term exposures which only measure effects on survival can be used to identify high levels of contamination, but may not be able to identify marginally contaminated sediments. The method described in this section can be used to evaluate potential effects of contaminated sediment on survival, growth, and reproduction of *H. azteca* in a 42-d test.

14.1.2 Section 14.2 describes general guidance for conducting a 42-d test with *H. azteca* that can be used to evaluate the effects of contaminants associated with sediments on survival, growth and reproduction. Refinements of these methods may be described in future editions of this manual after additional laboratories have successfully used the method (Section 17.6). The 42-d test with *H. azteca* has not been adequately evaluated in water with elevated salinity (Section 1.3.2).

14.1.3 The procedure outlined in Section 14.2 is based on procedures described in Ingersoll et al. (1998). The sediment exposure starts with 7- to 8-d-old amphipods. On Day 28, amphipods are isolated from the sediment and placed in water-only chambers where reproduction is measured on Day 35 and 42. Typically, amphipods are first in amplexus at about Day 21 to 28 with release of the first brood between Day 28 to 42. Endpoints measured include survival (Day 28, 35 and 42), growth (as length or dry weight measured on Day 28 and 42), and reproduction (number of young/female produced from Day 28 to 42). The procedures described in Table 14.1 include measurement of a variety of lethal and sublethal endpoints; minor modifications of the basic methods can be used in cases where only a subset of these endpoints is of interest.

14.1.3.1 Several designs were considered for measuring reproduction in sediment exposures based on the reproductive biology of *H. azteca* (Ingersoll et al., 1998). The first design considered was a continuation of the 28-d sediment exposures described in Ingersoll et al. (1996) for an additional two weeks to determine the number of young produced in the first brood. The limitation of this design is the difficulty in quantitatively isolating young amphipods from sediment (Tomasovic et al., 1995). A second design considered was extension of the 28-d sediment exposure for an additional month or longer until several broods are released. These multiple broods could then be isolated from the sediment. The limitation of this second design is that specific effects on reproduction could not be differentiated from reduced survival of offspring and it would still be difficult to isolate the young amphipods from sediment. A third design considered, and the one described in this manual, was to expose amphipods in sediment until a few days before the release of the first brood. The amphipods could then be sieved from the sediment and held in water to determine the number of young produced (Ingersoll et al., 1998). This test design allows a quantitative measure of reproduction. One limitation to this design is that amphipods might recover from effects of sediment exposure during this holding period in clean water (Landrum and Scavia, 1983; Kane Driscoll et al., 1997); however, amphipods are exposed to sediment during critical developmental stages before release of the first brood in clean water.

14.1.4 The method has been used to evaluate a formulated sediment and field-collected sediments with low to moderate concentrations of contaminants (Ingersoll et al., 1998). Survival of amphipods in these sediments was typically >85% after the 28-d sediment exposures and the 14-d holding period in water to measure reproduction (Ingersoll et al., 1998). The method outlined in 14.2 has also been evaluated in round-robin testing with 8 to 12 laboratories (Section 17.6). After the 28-d sediment exposures in a control sediment (West Bearskin), survival was >80% for >88% of the laboratories; length was >3.2 mm/individual for >71% of the laboratories; and dry weight was >0.15 mg/individual for >66% of the laboratories. Reproduction from Day 28 to Day 42 was >2 young/female for >71% of the laboratories participating in the round-robin testing. Reproduction was more variable within and among laboratories; hence, more replicates might be

Table 14.1 Test Conditions for Conducting a 42-d Sediment Toxicity Test with *Hyalella azteca* (modified from USEPA 1994a and ASTM 1999a).

	Parameter	Conditions
1.	Test type:	Whole-sediment toxicity test with renewal of overlying water
2.	Temperature:	23 ± 1°C
3.	Light quality:	Wide-spectrum fluorescent lights
4.	Illuminance:	About 100 to 1000 lux
5.	Photoperiod:	16L:8D
6.	Test chamber:	300-mL high-form lipless beaker
7.	Sediment volume:	100 mL
8.	Overlying water volume:	175 mL in the sediment exposure from Day 0 to Day 28 (175 to 275 mL in the water-only exposure from Day 28 to Day 42)
9.	Renewal of overlying water:	2 volume additions/d (Appendix A); continuous or intermittent (e.g., one volume addition every 12 h)
10.	Age of organisms:	7- to 8-d old at the start of the test
11.	Number of organisms/chamber:	10
12.	Number of replicate chambers/treatment:	12 (4 for 28-d survival and growth and 8 for 35- and 42-d survival, growth, and reproduction). Reproduction is more variable than growth or survival; hence, more replicates might be needed to establish statistical differences among treatments (See Section 14.2.3).
13.	Feeding:	YCT food, fed 1.0 mL (1800 mg/L stock) daily to each test chamber.
14.	Aeration:	None, unless dissolved oxygen in overlying water drops below 2.5 mg/L.
15.	Overlying water:	Culture water, well water, surface water or site water. Use of reconstituted water is not recommended.
16.	Test chamber cleaning:	If screens become clogged during a test, gently brush the *outside* of the screen (Appendix A).
17.	Overlying water quality:	Hardness, alkalinity, conductivity, and ammonia at the beginning and end of a sediment exposure (Day 0 and 28). Temperature daily. Conductivity weekly. Dissolved oxygen (DO) and pH three times/ week. Concentrations of DO should be measured more often if DO drops more than 1 mg/L since the previous measurement.
18.	Test duration:	42 d
19.	Endpoints:	28-d survival and growth; 35-d survival and reproduction; and 42-d survival, growth, reproduction, and number of adult males and females on Day 42
20.	Test acceptability:	Minimum mean control survival of 80% on Day 28. Additional performance-based criteria specifications are outlined in Table 14.3 based on results of round-robin testing (Sections 14.1.4 and 17.6).

needed to establish statistical differences among treatments with this endpoint.

14.1.5 Growth of *H. azteca* in sediment tests often provides unique information that can be used to discriminate toxic effects of exposure to contaminants (Brasher and Ogle, 1993; Borgmann, 1994; Kemble et al., 1994; Ingersoll et al., 1996; Kubitz et al., 1996; Milani et al., 1996; Steevens and Benson, 1998). Either length or weight can be measured in sediment tests with *H. azteca*. However, additional statistical options are available if length is measured on individual amphipods, such as nested analysis of variance which can account for variance in length between replicates (Steevens and Benson, 1998). Ongoing water-only studies testing select contaminants will provide additional data on the relative sensitivity and variability of sublethal endpoints in toxicity tests with *H. azteca* (Ingersoll et al., 1998).

14.1.6 Results of tests using procedures different from the procedures described in Section 14.2 may not be comparable, and these different procedures may alter contaminant bioavailability. Comparisons of results obtained using modified versions of these procedures might provide useful information concerning new concepts and procedures for conducting sediment tests with aquatic organisms. If tests are conducted with procedures different from the procedures described in this manual, additional tests are required to determine comparability of results (Section 1.3).

14.2 Procedure for Conducting a *Hyalella azteca* 42-d Test for Measuring the Effects of Sediment-associated Contaminants on Survival, Growth, and Reproduction

14.2.1 Conditions for evaluating sublethal endpoints in a sediment toxicity test with *H. azteca* are summarized in Table 14.1. A general activity schedule is outlined in Table 14.2. Decisions concerning the various aspects of experimental design, such as the number of treatments, number of test chambers/treatment, and water-quality characteristics should be based on the purpose of the test and the methods of data analysis (Section 16). When variability remains constant, the sensitivity of a test increases as the number of replicates increase.

14.2.2 The 42-d sediment toxicity test with *H. azteca* is conducted at 23°C with a 16L:8D photoperiod at an illuminance of about 100 to 1000 lux (Table 14.1). Test chambers are 300-mL high-form lipless beakers containing 100 mL of sediment and 175 mL of overlying water. Ten amphipods in each test chamber are fed 1.0 mL of YCT daily (Appendix B). Each test chamber receives 2 volume additions/d of overlying water. Water renewals may be manual or automated. Appendix A describes water-renewal systems that can be used to deliver overlying water. Overlying water should be a source of water that has been demonstrated to support survival, growth, and reproduction of *H. azteca* in culture. McNulty et al. (1999) and Kemble et al. (1999) observed poor survival of *H. azteca* in tests conducted 14 to 28 d using a variety of reconstituted waters including the reconstituted water (reformulated moderately hard reconstituted water) described in Smith et al. (1997) and described in the first edition of this manual (USEPA, 1994a). Borgmann (1996) described a reconstituted water that was used successfully to maintain *H. azteca* in culture; however, some laboratories have not had success when using this reconstituted

Table 14.2 General Activity Schedule for Conducting a 42-d Sediment Toxicity Test with *Hyalella azteca*

Day	Activity
Pre-Test	
-8	Separate known-age amphipods from the cultures and place in holding chambers. Begin preparing food for the test. The <24-h amphipods are fed 10 mL of YCT (1800 mg/L stock solution) and 10 mL of *Selenastrum capricornutum* (about 3.0 x 10^7 cells/mL) on the first day of isolation and 5 mL of both YCT and *S. capricornutum* on the 3rd and 5th d after isolation.
-7	Remove adults and isolate <24-h-old amphipods (if procedures outlined in Section 10.3.4 are followed).
-6 to -2	Feed and observe isolated amphipods (Section 10.3), monitor water quality (e.g., temperature and dissolved oxygen).
-1	Feed and observe isolated amphipods (Section 10.3), monitor water quality. Add sediment into each test chamber, place chambers into exposure system, and start renewing overlying water.
Sediment Test	
0	Measure total water quality (pH, temperature, dissolved oxygen, hardness, alkalinity, conductivity, ammonia). Transfer ten 7- to 8-d-old amphipods into each test chamber. Release organisms under the surface of the water. Add 1.0 mL of YCT (1800 mg/L stock) into each test chamber. Archive 20 test organisms for length determination or archive 80 test organisms for dry weight determination. Observe behavior of test organisms.
1 to 27	Add 1.0 mL of YCT to each test beaker. Measure temperature daily, conductivity weekly, and dissolved oxygen (DO) and pH three times/week. Observe behavior of test organisms.
28	Measure temperature, dissolved oxygen, pH, hardness, alkalinity, conductivity and ammonia. End the sediment-exposure portion of the test by collecting the amphipods with a #40-mesh sieve (425-µm mesh; U.S. standard size sieve). Use four replicates for growth measurements: count survivors and preserve organisms in sugar formalin for growth measurements. Use eight replicates for reproduction measurements: place survivors in individual replicate water-only beakers and add 1.0 mL of YCT to each test beaker/d and 2 volume additions/d (Appendix A) of overlying water.
Reproduction Phase	
29 to 35	Feed daily (1.0 mL of YCT). Measure temperature daily, conductivity weekly, and DO and pH three times a week. Measure hardness and alkalinity weekly. Observe behavior of test organisms.
35	Record the number of surviving adults and remove offspring. Return adults to their original individual beakers and add food.
36 to 41	Feed daily (1.0 mL of YCT). Measure temperature daily, conductivity weekly, and DO and pH three times a week. Measure hardness and alkalinity weekly. Observe behavior of test organisms.
41	Measure total water quality (pH, temperature, dissolved oxygen, hardness, alkalinity, conductivity, ammonia).
42	Record the number of surviving adults and offspring. Surviving adult amphipods on Day 42 are preserved in sugar formalin solution. The number of adult males in each beaker is determined from this archived sample. This information is used to calculate the number of young produced per female per replicate from Day 28 to Day 42.

water in the 42-d test (T.J. Norberg-King, USEPA, Duluth, MN, personal communication). For site-specific evaluations, the characteristics of the overlying water should be as similar as possible to the site where sediment is collected. Requirements for test acceptability are summarized in Table 14.3.

14.2.3 The number of replicates and concentrations tested depends in part on the significance level selected and the type of statistical analysis. A total of 12 replicates, each containing ten 7- to 8-d-old amphipods, are tested for each treatment. Starting the test with substantially younger or older organisms may compromise the reproductive endpoint. For the total of 12 replicates the assignment of beakers is as follows: 12 replicates are set up on Day -1 of which 4 replicates are used for 28-d growth and survival endpoints and the other 8 replicates are used for measurement of survival and reproduction on Day 35 and for measurement of survival, reproduction, or growth on Day 42.

14.3 General Procedures

14.3.1 Sediment into Test Chambers

14.3.1.1 The day before the sediment test is started (Day -1) each sediment should be thoroughly homogenized and added to the test chambers (Section 8.3.1). Sediment should be visually inspected to judge the degree of homogeneity. Excess water on the surface of the sediment can indicate separation of solid and liquid components. If a quantitative measure of homogeneity is

Table 14.3 Test Acceptability Requirements for a 42-d Sediment Toxicity Test with *Hyalella azteca*

A. It is recommended for conducting the 42-d test with *H. azteca* that the following performance criteria be met:

1. Age of *H. azteca* at the start of the test should be 7- to 8-d old. Starting a test with substantially younger or older organisms may compromise the reproductive endpoint.
2. Average survival of *H. azteca* in the control sediment on Day 28 should be greater than or equal to 80%.
3. Laboratories participating in round-robin testing (Section 17.6) reported after 28-d sediment exposures in a control sediment (West Bearskin), survival >80% for >88% of the laboratories; length >3.2 mm/individual for >71% of the laboratories; and dry weight >0.15 mg/individual for >66% of the laboratories. Reproduction from Day 28 to Day 42 was >2 young/female for >71% of the laboratories participating in the round-robin testing. Reproduction was more variable within and among laboratories; hence, more replicates might be needed to establish statistical differences among treatments with this endpoint.
4. Hardness, alkalinity, and ammonia in the overlying water typically should not vary by more than 50% during the sediment exposure, and dissolved oxygen should be maintained above 2.5 mg/L in the overlying water.

B. Performance-based criteria for culturing *H. azteca* include the following:

1. It may be desirable for laboratories to periodically perform 96-h water-only reference-toxicity tests to assess the sensitivity of culture organisms (Section 9.16.2). Data from these reference-toxicity tests could be used to assess genetic strain or life-stage sensitivity of test organisms to select chemicals.
2. Laboratories should track parental survival in the cultures and record this information using control charts if known-age cultures are maintained. Records should also be kept on the frequency of restarting cultures and the age of brood organisms.
3. Laboratories should record the following water-quality characteristics of the cultures at least quarterly: pH, hardness, alkalinity, and ammonia. Dissolved oxygen in the cultures should be measured weekly. Temperature of the cultures should be recorded daily. If static cultures are used, it may be desirable to measure water quality more frequently.
4. Laboratories should characterize and monitor background contamination and nutrient quality of food if problems are observed in culturing or testing organisms.
5. Physiological measurements such as lipid content might provide useful information regarding the health of the cultures.

C. Additional requirements:

1. All organisms in a test must be from the same source.
2. Storage of sediments collected from the field should follow guidance outlined in Section 8.2.
3. All test chambers (and compartments) should be identical and should contain the same amount of sediment and overlying water.
4. Negative-control sediment and appropriate solvent controls must be included in a test. The concentration of solvent used must not adversely affect test organisms.
5. Test organisms must be cultured and tested at 23°C (±1°C).
6. The mean of the daily test temperature must be within ±1°C of 23°C. The instantaneous temperature must always be within ±3°C of 23°C.
7. Natural physico-chemical characteristics of test sediment collected from the field should be within the tolerance limits of the test organisms.

required, replicate subsamples should be taken from the sediment batch and analyzed for TOC, chemical concentrations, and particle size.

14.3.1.2 Each test chamber should contain the same amount of sediment, determined either by volume or by weight. Overlying water is added to the chambers on Day -1 in a manner that minimizes suspension of sediment. This can be accomplished by gently pouring water along the sides of the chambers or by pouring water onto a baffle (e.g., a circular piece of Teflon with a handle attached) placed above the sediment to dissipate the force of the water. Renewal of overlying water is started on Day -1. A test begins when the organisms are added to the test chambers (Day 0).

14.3.2 Renewal of Overlying Water

14.3.2.1 Renewal of overlying water is required during a test. At any particular time during a test, flow rates through any two test chambers should not differ by more than 10%. Hardness, alkalinity and ammonia concentrations in the water above the sediment, within a treatment, typically should not vary by more than 50% during the test. Mount and Brungs (1967) diluters have been modified for sediment testing, and other automated water-delivery systems have also been used (Maki, 1977; Ingersoll and Nelson, 1990; Benoit et al., 1993; Zumwalt et al., 1994; Brunson et al., 1998; Wall et al., 1998; Leppanen and Maier, 1998). The water-delivery system should be calibrated before a test is started to verify that the system is functioning properly. Renewal of overlying water is started on Day -1 before the addition of test organisms or food on Day 0. Appendix A describes water-renewal systems that can be used for conducting sediment tests.

14.3.2.2 In water-renewal tests with one to four volume additions of overlying water/d, water-quality characteristics generally remain similar to the inflowing water (Ingersoll and Nelson, 1990; Ankley et al., 1993); however, in static tests, water quality may change profoundly during the exposure (Shuba et al., 1978). For example, in static whole-sediment tests, the alkalinity, hardness, and conductivity of overlying water more than doubled in several treatments during a four-week exposure (Ingersoll and Nelson, 1990). Additionally, concentrations of metabolic products (e.g., ammonia) may also increase during static exposures, and these compounds can either be directly toxic to the test organisms or may contribute to the toxicity of the contaminants in the sediment. Furthermore, changes in water-quality characteristics such as hardness may influence the toxicity of many inorganic (Gauss et al., 1985) and organic (Mayer and Ellersieck, 1986) contaminants. Although contaminant concentrations are reduced in the overlying water in water-renewal tests, organisms in direct contact with sediment generally receive a substantial proportion of a contaminant dose directly from either the whole sediment or from the pore water.

14.3.3 Acclimation

14.3.3.1 Test organisms must be cultured and tested at 23°C. Ideally, test organisms should be cultured in the same water that will be used in testing. However, acclimation of test organisms to the test water is not required.

14.3.3.2 Culturing of organisms and toxicity assessment are typically conducted at 23°C. However, occasionally there is a need to perform evaluations at temperatures different than that recommended. Under these circumstances, it may be necessary to acclimate organisms to the desired test temperature to prevent thermal shock when moving immediately from the culture temperature to the test temperature (ASTM, 1999a). Acclimation can be achieved by exposing organisms to a gradual change in temperature; however, the rate of change should be relatively slow to prevent thermal shock. A change in temperature of 1°C every 1 to 2 h has been used successfully in some studies (P.K. Sibley, University of Guelph, Guelph, Ontario, personal communication; APHA, 1989). Testing at temperatures other than 23°C needs to be preceded by studies to determine expected performance under alternate conditions.

14.3.4 Placing Organisms in Test Chambers

14.3.4.1 Test organisms should be handled as little as possible. Amphipods should be introduced into the overlying water below the air-water interface. Test organisms can be pipetted directly into overlying water. The size of the test organisms at the start of the test should be measured using the same measure (length or weight) that will be used to assess their size at the end of the test. For length, a minimum of 20 organisms should be measured. For weight measurement, a larger sample size (e.g., 80) may be desirable because of the relatively small mass of the organisms. This information can be used to determine consistency in the size of the organisms used to start a test.

14.3.5 Feeding

14.3.5.1 For each beaker, 1.0 mL of YCT is added from Day 0 to Day 42. Without addition of food, the test organisms may starve during exposures. However, the addition of the food may alter the availability of the contaminants in the sediment (Wiederholm et al., 1987; Harkey et al., 1994). Furthermore, if too much food is added to the test chamber, or if the mortality of test organisms is high, fungal or bacterial growth may develop on the sediment surface. Therefore, the amount of food added to the test chambers is kept to a minimum.

14.3.5.2 Suspensions of food should be thoroughly mixed before aliquots are taken. If excess food collects on the sediment, a fungal or bacterial growth may develop on the sediment surface, in which case feeding should be suspended for one or more days. A drop in dissolved oxygen below 2.5 mg/L during a test may indicate that the food added is not being consumed. Feeding should be suspended for the amount of time necessary to increase the

dissolved oxygen concentration (ASTM, 1999a). If feeding is suspended in one treatment, it should be suspended in all treatments. Detailed records of feeding rates and the appearance of the sediment surface should be made daily.

14.3.6 Monitoring a Test

14.3.6.1 All chambers should be checked daily and observations made to assess test organism behavior such as sediment avoidance. However, monitoring effects on burrowing activity of test organisms may be difficult because the test organisms are often not visible during the exposure. The operation of the exposure system should be monitored daily.

14.3.6.2 Measurement of Overlying Water-quality Characteristics

14.3.6.2.1 Conductivity, pH, DO, hardness, alkalinity, and ammonia should be measured in all treatments at the beginning and at the end of the sediment exposure portion of the test. Water-quality characteristics should also be measured at the beginning and end of the reproductive phase (Day 29 to Day 42). Conductivity should be measured weekly, whereas pH and DO should be measured three times/week (Section 14.3.6.2.2). Overlying water should be sampled just before water renewal from about 1 to 2 cm above the sediment surface using a pipet. It may be necessary to composite water samples from individual replicates. The pipet should be checked to make sure no organisms are removed during sampling of overlying water.

14.3.6.2.2 Dissolved oxygen should be measured three times/week and should be at a minimum of 2.5 mg/L. If a probe is used to measure dissolved oxygen in overlying water, it should be thoroughly inspected between samples to make sure that organisms are not attached and should be rinsed between samples to minimize cross contamination. Aeration can be used to maintain dissolved oxygen in the overlying water above 2.5 mg/L (i.e., about 1 bubble/second in the overlying water). Dissolved oxygen and pH can be measured directly in the overlying water with a probe.

14.3.6.2.3 Temperature should be measured at least daily in at least one test chamber from each treatment. The temperature of the water bath or the exposure chamber should be continuously monitored. The daily mean test temperature must be within ±1°C of 23°C. The instantaneous temperature must always be within ±3°C of 23°C.

14.3.7 Ending a Test

14.3.7.1 Endpoints monitored include 28-d survival and growth of amphipods and 35-d and 42-d survival, growth, and reproduction (number of young/female) of amphipods. Growth or reproduction of amphipods may be a more sensitive toxicity endpoint compared to survival (Burton and Ingersoll, 1994; Kemble et al., 1994; Ingersoll et al., 1998).

14.3.7.2 On Day 28, 4 of the replicate beakers/sediment are sieved with a #40-mesh sieve (425-µm mesh; U.S. standard size sieve) to remove surviving amphipods for growth determinations. Any of the surviving amphipods in the water column or on the surface of the sediment can be pipetted from the beaker before sieving the sediment. The sediment in each beaker should be sieved in two separate aliquots (i.e., most of the amphipods will probably be found in the surface aliquot). Immobile organisms isolated from the sediment surface or from sieved material should be considered dead. Surviving amphipods from these 4 replicates can be preserved in separate vials containing 8% sugar formalin solution if length of amphipods is to be measured (Ingersoll and Nelson, 1990). The sugar formalin solution is prepared by adding 120 g of sucrose to 80 mL of formalin which is then brought to a volume of 1 L using deionized water. This stock solution is mixed with an equal volume of deionized water when used to preserve organisms. NoTox® (Earth Safe Industries, Belle Mead, NJ) can be used as a substitute for formalin (Unger et al., 1993).

14.3.7.3 A consistent amount of time should be taken to examine sieved material for recovery of test organisms (e.g., 5 min/replicate). Laboratories should demonstrate that their personnel are able to recover an average of at least 90% of the organisms from whole sediment. For example, test organisms could be added to control or test sediments, and recovery could be determined after 1 h (Tomasovic et al., 1994).

14.3.7.4 Growth of amphipods can be reported as either length or weight; however, additional statistical options are available if length is measured on individual organisms (Section 14.4.5.3).

14.3.7.5 Amphipod body length (±0.1 mm) can be measured from the base of the first antenna to the tip of the third uropod along the curve of the dorsal surface (Figure 11.1). Kemble et al. (1994) describe the use of a digitizing system and microscope to measure lengths of *H. azteca*. Kemble et al. (1994) also photographed invertebrates (at a magnification of 3.5X) and measured length using a computer-interfaced digitizing tablet.

14.3.7.6 Dry weight of amphipods in each replicate can be determined on Day 28 and 42. If both weight and length are to be determined, weight should be measured after length on the preserved samples. Gaston et al. (1995) and Duke et al. (1996) have shown that biomass or length of several aquatic invertebrates did not significantly change after two to four weeks of storage in 10% formalin. If test organisms are to be used for an evaluation of bioaccumulation, it is not advisable to dry the sample before conducting the residue analysis. If conversion from wet weight to dry weight is necessary, aliquots of organisms can be weighed to establish wet to dry weight conversion factors. A consistent procedure should be used to remove the excess water from the organisms before measuring wet weight.

14.3.7.7 Dry weight of amphipods can be determined as follows: (1) transferring the archived amphipods from a replicate out of the sugar formalin solution into a crystallizing dish; (2) rinsing amphipods with deionized water; (3) transferring these rinsed amphipods to a preweighed aluminum pan; (4) drying these samples for 24 h at 60°C; and (5) weighing the pan and dried amphipods on a balance to the nearest 0.01 mg. Average dry weight of individual amphipods in each replicate is calculated from these data. Due to the small size of the amphipods, caution should be taken during weighing (10 dried amphipods after a 28-d sediment exposure may weigh less than 2.5 to 3.5 mg). Weigh pans need to be carefully handled using powder-less gloves and the balance should be calibrated with standard weights with each use. Use of small aluminum pans (e.g., 7 x 22 x 7 mm, Sigma Chemical Company, St. Louis, MO) will help reduce variability in measurements of dry weight. Weigh boats can also be constructed from sheets of aluminum foil.

14.3.7.8 The first edition of this manual (USEPA, 1994a) recommended dry weight as a measure of growth for both *H. azteca* and *C. tentans*. For *C. tentans*, this recommendation was changed in the current edition to ash-free dry weight (AFDW) instead of dry weight, with the intent of reducing bias introduced by gut contents (Sibley et al., 1997a). However, this recommendation was not extended to include *H. azteca*. Studies by Dawson et al. (personal communication, T.D. Dawson, Integrated Laboratory Systems, Duluth, MN) have indicated that the ash content of *H. azteca* is not greatly decreased by purging organisms in clean water before weighing, suggesting that sediment does not comprise a large portion of the overall dry weight. In addition, using AFDW further decreases an already small mass, potentially increasing measurement error. For this reason, dry weight continues to be the recommended endpoint for estimating growth of *H. azteca* via weight (growth can also be determined via length).

14.3.7.9 On Day 28, the remaining 8 beakers/sediment are also sieved and the surviving amphipods in each sediment beaker are placed in 300-mL water-only beakers containing 150 to 275 mL of overlying water and a 5-cm x 5-cm piece of Nitex screen (Nylon Bolting cloth; 44% open area and 280-um aperture, Wildlife Supply Company, Saginaw, MI; Ingersoll et al., 1998). In a subsequent study, improved reproduction of *H. azteca* was observed when the Nitex screen was replaced with a 3-cm x 3-cm piece of the nylon "Coiled-web material" described in Section 10.3.4 for use in culturing amphipods (T.J. Norberg-King, USEPA, personal communication). Each water-only beaker receives 1.0 mL of YCT stock solution and about two volume additions of water daily.

14.3.7.10 Reproduction of amphipods is measured on Day 35 and Day 42 in the water-only beakers by removing and counting the adults and young in each beaker. On Day 35, the adults are then returned to the same water-only beakers. Adult amphipods surviving on Day 42 are preserved in sugar formalin. The number of adult females is determined by simply counting the adult males (mature male amphipods will have an enlarged second gnathopod) and assuming all other adults are females (cf., Figure 11.1). The number of females is used to determine number of young/female/beaker from Day 28 to Day 42. Growth can also be measured for these adult amphipods.

14.4 Interpretation of Results

14.4.1 Data Analysis

14.4.1.1 Endpoints measured in the 42-d *H. azteca* test include survival (Day 28, 35, and 42), growth (as length or dry weight on Day 28 and 42), and reproduction (number of young/female produced from Day 28 to 42). Section 16 describes general information regarding statistical analysis of these data, including both point estimates (i.e., LC50s) and hypothesis testing (i.e., ANOVA). The following sections describe species-specific information that is useful in helping to interpret the results of 42-d sediment toxicity tests with *H. azteca.*

14.4.2 Age Sensitivity

14.4.2.1 The sensitivity of *H. azteca* appears to be relatively similar up to at least 24- to 26-d-old organisms (Collyard et al., 1994). For example, the toxicity of diazinon, Cu, Cd, and Zn was similar in 96-h water-only exposures starting with 0- to 2-d-old organisms through 24- to 26--d-old organisms (Figure 11.2). The toxicity of alkylphenol ethoxylate (a surfactant) tended to increase with age. In general, this suggests that tests started with 7-d to 8-d-old amphipods would be representative of the sensitivity of *H. azteca* up to at least the adult life stage.

14.4.3 Grain Size

14.4.3.1 *Hyalella azteca* tolerate a wide range in sediment grain size and organic matter in 10- to 28-d tests measuring effects on survival or growth (Ankley et al., 1994; Suedel and Rodgers, 1994; Ingersoll et al., 1996; Kemble et al., 1999). Using the method outlined in Section 14.2, no significant correlations were observed between the survival, growth, or reproduction of *H. azteca* and the physical characteristics of the sediment (grain size ranging from predominantly silt to predominantly sand), TOC (ranging from 0.3 to 9.6%), water content (ranging from 19 to 81%; Ingersoll et al., 1998). Additionally, no significant correlations were observed between these biological endpoints and the water-quality characteristics (i.e., hardness, alkalinity, ammonia) of pore water or overlying water in the sediments evaluated by Ingersoll et al. (1998). Weak trends were observed between reproduction of amphipods and percent clay, percent silt, and percent sand. Additional study is needed to better evaluate potential relationships between reproduction of *H. azteca* and these physical characteristics of the sediment. The weak relationship between the sediment grain size and reproduction may have been due to the fact that samples with higher amounts of sand also had higher concentrations of organic contaminants compared to other samples evaluated in Ingersoll et al. (1998).

14.4.3.2 Until additional studies have been conducted which substantiate this lack of a correlation between physical characteristics of sediment and the reproductive endpoints measured in the long-term sediment test with *H. azteca*, it would be desirable to test control or reference sediments which are representative of the physical characteristics of field-collected sediments. Formulated sediments could be used to bracket the ranges in physical characteristics expected in the field-collected sediments being evaluated (Section 7.2). Addition of YCT should provide a minimum amount of food needed to support adequate survival, growth, and reproduction of *H. azteca* in sediments low in organic matter. Without addition of food, *H. azteca* can starve during exposures (McNulty et al., 1999) making it impossible to differentiate effects of contaminants from other sediment characteristics.

14.4.4 Influence of Indigenous Organisms

14.4.4.1 Survival of *H. azteca* in 28-d tests was not reduced in the presence of oligochaetes in sediment samples (Reynoldson et al., 1994). However, growth of amphipods was reduced when high numbers of oligochaetes were placed in a sample. Therefore, it is important to determine the number and biomass of indigenous organisms in field-collected sediments in order to better interpret growth data (Reynoldson et al., 1994; DeFoe and Ankley, 1998). Furthermore, presence of predators may also influence response of test organisms in sediment (Ingersoll and Nelson, 1990).

14.4.5 Relationships between Growth and Reproductive Endpoints

14.4.5.1 Natural or anthropogenic stressors that affect growth of invertebrates may also affect reproduction, because of a minimum size needed for reproduction (Rees and Crawley, 1989; Ernsting et al., 1993; Moore and Dillon, 1993; Enserink et al., 1995; Moore and Farrar, 1996; Sibley et al., 1996, 1997a). Ingersoll et al. (1998) reported a significant correlation between reproduction from Day 28 to 42 and length of *H. azteca* on Day 28 when data are plotted by the mean of each treatment (Figure 14.1a; Spearman rank correlation of 0.59, p=0.0001). Based on 28-d lengths, smaller amphipods (<3.5 mm) tended to have lower reproduction and larger amphipods (>4.3 mm) tended to have higher reproduction; however, the range in reproduction was wide for amphipods 3.5 to 4.3 mm in length. Based on 42-d lengths, there was a weaker correlation between length and reproduction (i.e., reproduction and length measured in paired replicates; Figure 14.1b, Spearman rank correlation of 0.49, p=0.0001). Similarly, plotting data by individual replicates (data not shown) did not improve the relationship between 42-d length and reproduction compared to the plots by the mean of each treatment (Figure 14.1b; Ingersoll et al., 1998).

14.4.5.2 Weaker relationships were observed between reproduction and dry weight measured on Day 28 (Figure 14.2a, Spearman rank correlation of 0.44, p = 0.0037, n = 42) or dry weight measured on Day 42 (Figure 14.2b, Spearman rank correlation 0.34, p = 0.0262, n = 42). Round-robin studies (Section 17.6) have generated additional data that will be used to further evaluate relationships between growth and reproduction of *H. azteca* in sediment tests using the procedures outlined in Section 14.2.

14.4.5.3 A significant correlation was evident between length and dry weight of amphipods (Figure 14.3, Spearman rank of 0.80, p=0.0001) indicating that either length or weight could be measured in sediment tests with *H. azteca*. However, additional statistical options are available if length is measured on individual amphipods, such as nested ANOVA which can account for variance in length within replicates (Steevens and Benson, 1998). Analyses are ongoing to evaluate the ability of length vs. weight to discriminate between contaminated and uncontaminated samples in a database described in Ingersoll et al. (1996).

14.4.5.4 The relatively variable relationship between growth and reproduction probably reflects the fact that most of these comparisons were made within a fairly narrow range in length (3.5 to 5.0 mm; Figure 14.1) or dry weight (0.25 to 0.50 mg; Figure 14.2). Other investigators have reported a similar degree of variability in reproduction of *H. azteca* within a narrow range of length or weight, with stronger correlations observed over wider ranges (Hargrave, 1970b; Strong, 1972; Wen, 1993; Moore and Farrar, 1996). The degree of correlation between growth and reproduction may also be dependent on the genetic strain of *H. azteca* evaluated (Strong, 1972; France, 1992).

14.4.5.5 The proportion of males to females within a treatment or by replicate was not correlated to young production, but may have contributed to a variation in reproduction (Ingersoll et al., 1998). Wen (1993) reported that when two or three males were placed in a beaker with one female *H. azteca*, the frequency of successful amplexus was reduced, possibly from aggression between the males. Future study is needed to determine if increasing the number of amphipods/beaker would result in a more consistent proportion of males to females within a beaker and would reduce variability in reproduction.

14.4.5.6 Reproduction was often more variable than growth (Ingersoll et al., 1998). The coefficient of variation (CV) was typically <10% for growth and >20% for reproduction. This difference in variation affects the statistical power of the comparisons and the number of replicates required for a test. For example, detection of a 20% difference between treatment means at a statistical power of 0.8 would require about 4 replicates at a CV of 10% and 14 replicates at a CV of 20% (Figure 16.5). Fewer replicates would be required if detection of larger differences among treatment means were of interest. Ongoing water-only studies testing select contaminants will hopefully provide additional data on the relative sensitivity and variability of sublethal endpoints in toxicity tests with *H. azteca* (Ingersoll et al., 1998).

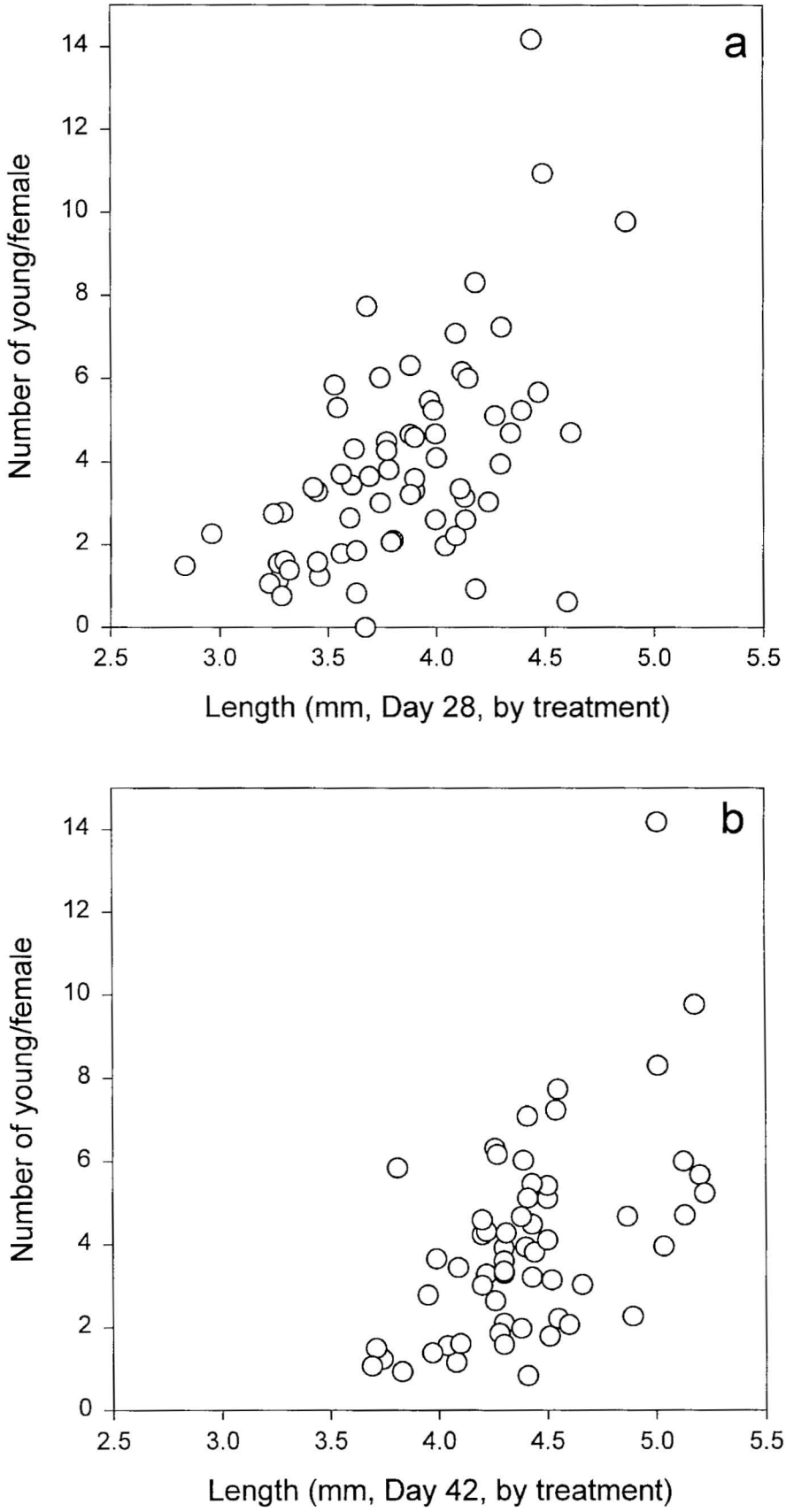
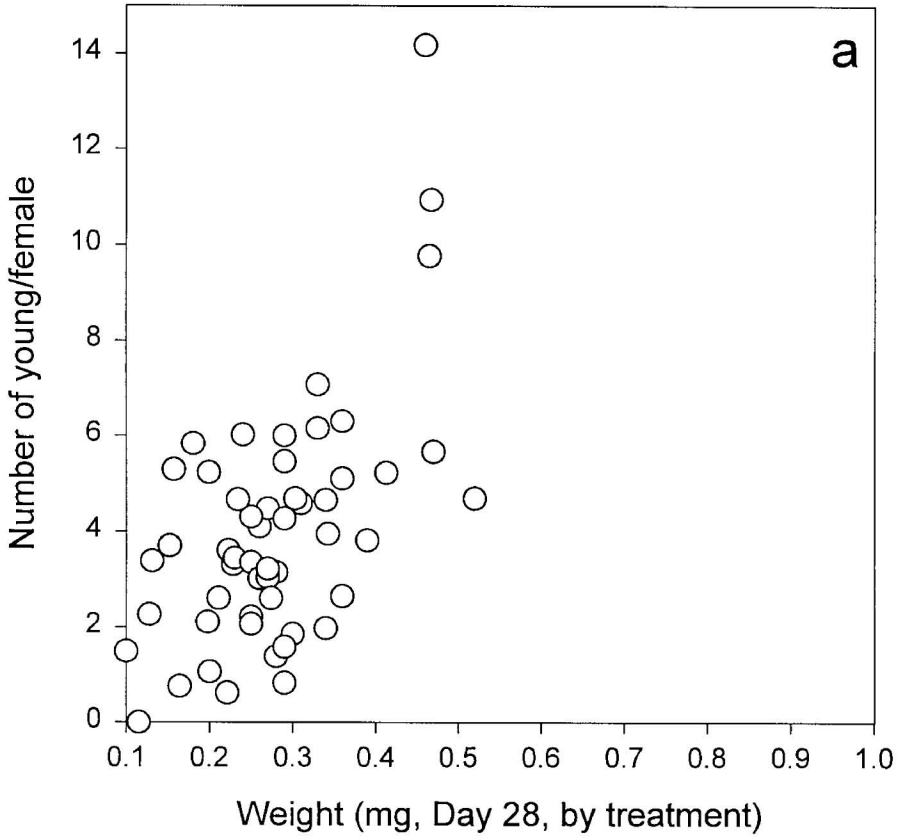


Figure 14.1 **Relationships between *Hyalella azteca* length and reproduction by (a) treatment means for 28-d length or (b) treatment means for 42-d length.**



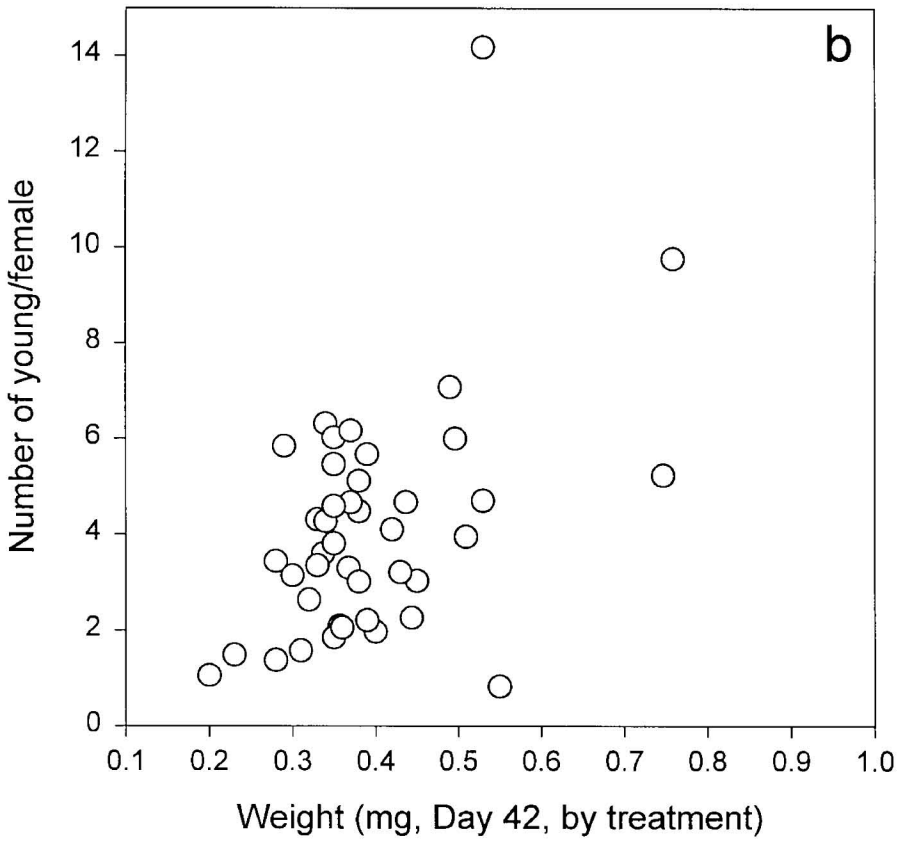


Figure 14.2 Relationships between *Hyalella azteca* dry weight and reproduction by (a) treatment means for 28-d dry weight or (b) treatment means for 42-d dry weight.

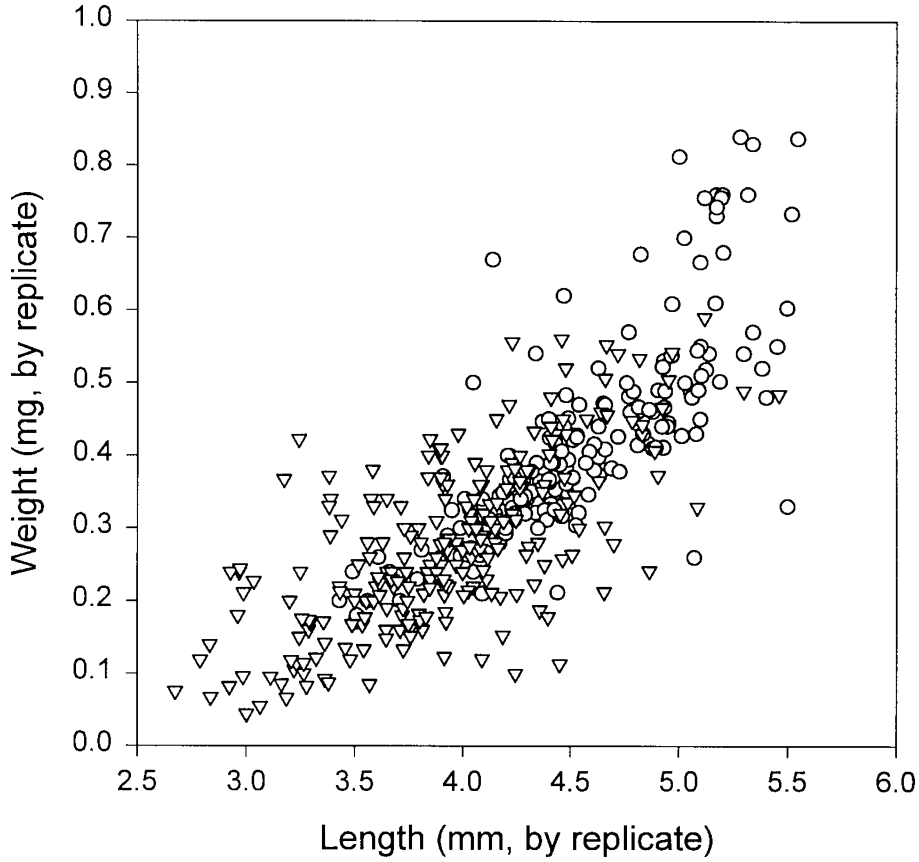


Figure 14.3 Relationship between *Hyalella azteca* length and dry weight. Triangles are data for Day 28 and circles are data for Day 42 (Ingersoll et al., 1998).

14.4.5.7 The 8-replicate design recommended in this manual (Table 14.1) is a compromise between logistical constraints and statistical considerations. Laboratories experienced with this method have shown CVs of 25 to 50% (Ingersoll et al., 1998), though some higher values were observed during the round-robin testing (Section 17.6), in which most labs had not previously performed the test.

14.4.5.8 As discussed above, the number of replicates can be adjusted according to the needs of a particular study. For example, Kubitz et al. (1996) recommended a two-step process for assessing growth in sediment tests with *H. azteca*. Using this process, a limited number of replicates would be tested in a screening step. Samples identified as possibly affecting reproduction could then be tested in a confirmatory step with additional replicates. This two-step analysis conserves laboratory resources and increases statistical power when needed to discriminate sublethal effects. A similar approach could be applied to evaluate reproductive effects of contaminants in sediment where a limited number of replicates could be initially tested to evaluate potential effects. Samples identified as possibly toxic based on reproduction could then be reevaluated using an increased number of replicates. However, the use of sediments stored for extended periods of time may introduce variability in results between the two studies (Section 8.2).

14.4.6 Relative Endpoint Sensitivity

14.4.6.1 Measurement of sublethal endpoints in sediment tests with *H. azteca* can provide unique information that has been used to discriminate toxic effects of exposure to contaminants. Table 14.4 compares the relative sensitivity of survival and growth endpoints in 14- and 28-d tests with *H. azteca* (Ingersoll et al., 1996, 1998). When 14-d and 28-d tests were conducted concurrently measuring both survival and growth, both tests identified 34% of the samples as toxic and 53% of the samples as not toxic (N=32). Both tests identified an additional 6% of the samples as toxic. Survival or growth endpoints identified a similar percentage of samples as toxic in both the 14- and 28-d tests. However, the majority of the samples used to make these comparisons were highly contaminated. Additional exposures conducted with moderately contaminated sediment might exhibit a higher percentage of sublethal effects in the 28-d test compared to the 14-d test.

14.4.6.2 When both survival and growth were measured in 14-d tests (N=25), only 4% of the samples reduced

Table 14.4 Percentage of Paired Tests or Paired Endpoints Identifying Samples as Toxic in *Hyalella azteca* 14-d or 28-d Tests. See USEPA (1996a) and Ingersoll et al. (1996) for a description of this database.

Comparisons	Tox/tox[1] (%)	Not/not[2] (%)	Tox/not[3] (%)	Not/tox[4] (%)	N[5]
Survival or growth: 14 d/28 d	34	53	6	6	32
Survival: 14 d/28 d	25	66	0	10	32
Growth: 14 d/28 d	8	64	12	16	25
14 d: survival/growth	4	60	20	16	25
28 d: survival/growth	16	52	14	18	44

[1] Tox/tox: samples toxic (significant reduction relative to the control $p<0.05$) with both tests (or both endpoints).
[2] Not/not: samples not toxic with both tests (or both endpoints).
[3] Tox/not: samples toxic to the first but not the second test (or endpoint).
[4] Not/tox: samples not toxic to the first but toxic to the second test (or endpoint).
[5] N: number of samples

both survival and growth; however, 20% reduced survival only and 16% reduced growth only (60% did not reduce survival or growth). Hence, if survival was the only endpoint measured in 14-d tests, 16% of the toxic samples would be incorrectly classified. Similar percentages are also observed for the 28-d tests. When both survival and growth were measured in the 28-d test (N=44), 16% of the samples reduced both survival and growth, 14% reduced survival only, 18% reduced growth only, and 52% did not reduce survival or growth.

14.4.6.3 The endpoint comparisons in Table 14.4 represent only samples where both survival and growth could be measured. If a sample was extremely toxic, it would not be included in this comparison since growth could not be measured. Moderately contaminated sediments that did not severely reduce survival could have a reduced growth. For example, in 28-d tests with sediments from the Clark Fork River, growth was a more sensitive endpoint compared to survival or maturation. Only 13% of the samples reduced survival and 20% of the samples reduced maturation; however, growth was reduced in 53% of the samples (Kemble et al., 1994).

14.4.6.4 Other investigators have reported measurement of growth in tests with *H. azteca* often provides unique information that can help discriminate toxic effects of exposure to contaminants in sediment (Kubitz et al., Milani et al., 1996; Steevens and Benson, 1998) or water (Brasher and Ogle, 1993; Borgmann, 1994). Similarly, in sediment tests with the midge *C. tentans*, sublethal endpoints are often more sensitive than survival as indicators of contaminant stress (Section 12 and 15). In contrast, Borgmann et al. (1989) reported that growth or reproduction did not add additional information beyond measurement of survival of *H. azteca* in water-only exposures with cadmium or pentachlorophenol. Similarly, Day et al. (1995) reported that weight did not add additional information beyond measurement of survival in 28-d tests with *H. azteca*. Ramirez-Romero (1997) reported that reproduction of *H. azteca* was not affected by exposure to sublethal concentrations of fluoranthene in sediment when exposures were started with juvenile amphipods. Brasher and Ogle (1993) started exposures with adult amphipods and observed the sensitivity of reproduction compared to survival of *H. azteca* was dependent on the chemical tested (reproduction more sensitive to selenite and survival more sensitive to selenate in water-only exposures). Long-term exposures starting with juvenile amphipods would likely be more appropriate to assess effects of contaminants on reproduction (i.e., Carr and Chapman, 1992; Nebeker et al., 1992).

14.4.7 Future Research

14.4.7.1 Additional studies are needed to further evaluate the use of reconstituted water and ammonia on long-term exposures with *H. azteca*. Section 1.3.8.5 addresses interpretative guidance for evaluating toxicity associated with ammonia in sediment. Ongoing water-only toxicity tests with select chemicals (i.e., cadmium, DDD and fluoranthene) should generate data that can be used to better determine the relative sensitivity of survival, reproduction, and growth endpoints in tests with *H. azteca* (Ingersoll et al., 1998). These water-only studies will also be used to evaluate potential recovery of amphipods after transfer into clean water to measure reproduction. In addition to studies evaluating the relative sensitivity of endpoints, research is also needed to evaluate the ability of these laboratory endpoints to estimate responses of benthic organisms exposed in the field to chemicals in sediments (Canfield et al., 1996).

Section 15
Test Method 100.5
Life-cycle Test for Measuring the Effects of Sediment-associated Contaminants on *Chironomus tentans*

15.1 Introduction

15.1.1 The midge *Chironomus tentans* has been used extensively in the short-term assessment of chemicals in sediments (Wentsel et al., 1977; Nebeker et al., 1984; Giesy et al., 1988; West et al., 1994), and standard methods have been developed for testing with this midge using 10-d exposures (Ingersoll et al., 1995; USEPA, 1994a; ASTM, 1999a). *Chironomus tentans* is a good candidate for long-term toxicity testing because it normally completes its life cycle in a relatively short period of time (25 to 30 d at 23°C), and a variety of developmental (growth, survivorship) and reproductive (fecundity) endpoints can be monitored. In addition, emergent adults can be readily collected so it is possible to transfer organisms from the sediment test system to clean, overlying water for direct quantification of reproductive success.

15.1.2 The long-term sediment toxicity test with the midge, *Chironomus tentans*, is a life-cycle test in which the effects of sediment exposure on survival, growth, emergence, and reproduction are assessed (Benoit et al., 1997). Procedures for conducting the long-term test with *C. tentans* are described in Section 15.2. The test is started with newly hatched larvae (<24-h old) and continues through emergence, reproduction, and hatching of the F_1 generation. Survival is determined at 20 d and at the end of the test (about 50 to 65 d). Growth is determined at 20 d, which corresponds to the 10-d endpoint in the 10-d *C. tentans* growth test started with 10-d-old larvae (Section 12). From Day 23 to the end of the test, emergence and reproduction are monitored daily. The number of eggs is determined for each egg case, which is incubated for 6 d to determine hatching success. Each treatment of the life-cycle test is ended separately when no additional emergence has been recorded for 7 consecutive days (the 7-d criterion). When no emergence is recorded from a treatment, ending of that treatment should be based on the control sediment using this 7-d criterion. Appendix C and Table 6.1 outline equipment and supplies needed to conduct this test. The procedures described in Table 15.1 include measurement of a variety of lethal and sublethal endpoints; minor modifications of the basic methods can be used in cases where only a subset of these endpoints is of interest.

15.1.3 The method outlined in Section 15.2 has been evaluated in round-robin testing with 10 laboratories using two clean sediments (Section 17.6). In the preliminary round-robin with 1.5 mL of Tetrafin/d as a food source, 90% of labs met the survival criterion (>70%), 100% of labs met the growth criterion (>0.48 mg AFDW), 70% of labs met the emergence criterion (>50%), 90% of labs met the reproduction criterion (>800 eggs/female), and 88% of labs met the percent hatch criterion (>80%). Reproduction was generally more variable than growth or survival within and among laboratories; hence, more replicates might be needed to establish statistical significance of small decreases in reproduction.

15.1.4 Growth and other sublethal endpoints in sediment tests with *C. tentans* often provide unique information that can be used to discriminate toxic effects of exposure to contaminants. See Section 15.4.6 for additional details.

15.1.5 Results of tests using procedures different from the procedures described in Section 15.2 may not be comparable and these different procedures may alter contaminant bioavailability. Comparison of results obtained using modified versions of these procedures might provide useful information concerning new concepts and procedures for conducting sediment tests with aquatic organisms. If tests are conducted with procedures different from the procedures described in this manual, additional tests are required to determine comparability of results (Section 1.3).

15.2 Procedure for Conducting a Life-cycle Test for Measuring the Effects of Sediment-associated Contaminants on *Chironomus tentans*

15.2.1 Conditions for conducting a long-term sediment toxicity test with *C. tentans* are summarized in Table 15.1. A general activity schedule is outlined in Table 15.2. Decisions concerning the various aspects of experimental design, such as the number of treatments, number of test chambers/treatment, and water-quality characteristics should be based on the purpose of the test and the methods of data analysis (Section 16). When variability

Table 15.1 Test Conditions for Conducting a Long-term Sediment Toxicity Test with *Chironomus tentans*

	Parameter	Conditions
1.	Test type:	Whole-sediment toxicity test with renewal of overlying water
2.	Temperature:	23 ±1°C
3.	Light quality:	Wide-spectrum fluorescent lights
4.	Illuminance:	About 100 to 1000 lux
5.	Photoperiod:	16L:8D
6.	Test chamber:	300-mL high-form lipless beaker
7.	Sediment volume:	100 mL
8.	Overlying water volume:	175 mL
9.	Renewal of overlying water:	2 volume additions/d (Appendix A); continuous or intermittent (e.g., one volume addition every 12 h)
10.	Age of organisms:	< 24-h-old larvae
11.	Number of organisms/chamber:	12
12.	Number of replicate chambers/treatment:	16 (12 at Day -1 and 4 for auxiliary males on Day 10)
13.	Feeding:	Tetrafin® goldfish food, fed 1.5 mL daily to each test chamber starting Day -1 (1.0 mL contains 4.0 mg of dry solids)
14.	Aeration:	None, unless dissolved oxygen in overlying water drops below 2.5 mg/L
15.	Overlying water:	Culture water, well water, surface water, site water, or reconstituted water
16.	Test chamber cleaning:	If screens become clogged during a test, gently brush the *outside* of the screen (Appendix A).
17.	Overlying water quality:	Hardness, alkalinity, conductivity, and ammonia at the beginning, on Day 20, and at the end of a test. Temperature daily (ideally continuously). Dissolved oxygen (DO) and pH three times/week. Conductivity weekly. Concentrations of DO should be measured more often if DO has declined by more than 1 mg/L since previous measurement.
18.	Test duration:	About 50 to 65 d; each treatment is ended separately when no additional emergence has been recorded for seven consecutive days. When no emergence is recorded from a treatment, termination of that treatment should be based on the control sediment using this 7-d criterion.
19.	Endpoints:	20-d survival and weight; female and male emergence, adult mortality, the number of egg cases oviposited, the number of eggs produced, and the number of hatched eggs. Potential sublethal endpoints are listed in Table 15.4.
20.	Test acceptability:	Average size of *C. tentans* in the control sediment at 20 d must be at least 0.6 mg/surviving organism as dry weight or 0.48 mg/surviving organism as AFDW. Emergence should be greater than or equal to 50%. Experience has shown that pupae survival is typically >83% and adult survival is >96%. Time to death after emergence is <6.5 d for males and <5.1 d for females. The mean number of eggs/egg case should be greater than or equal to 800 and the percent hatch should be greater than or equal to 80%. See Sections 15.1.3 and 17.6 for a summary of performance in round-robin testing.

remains constant, the sensitivity of a test increases as the number of replicates increases.

15.2.2 The long-term sediment toxicity test with *C. tentans* is conducted at 23°C with a 16L:8D photoperiod at an illuminance of about 100 to 1000 lux (Table 15.1). Test chambers are 300-mL high-form lipless beakers containing 100 mL of sediment and 175 mL of overlying water. Each test chamber receives 2 volume additions/d of overlying water. Water renewals may be manual or automated. Appendix A describes water-renewal systems that can be used to deliver overlying water. Overlying water should be a source of water that has been demonstrated to support survival, growth, and reproduction of *C. tentans* in culture. For site-specific evaluations, the characteristics of the overlying water should be as similar as possible to the site where sediment is collected. Requirements for test acceptability are summarized in Table 15.3.

15.2.3 The number of replicates and concentrations tested depends in part on the significance level selected and the type of statistical analysis. For routine testing, a total of 16 replicates, each containing 12, <24-h-old larvae are tested for each treatment. For the total of 16 replicates the assignment of beakers is as follows: initially, 12 replicates are set up on Day -1 of which 4 replicates are used for 20-d growth and survival endpoints and 8

Table 15.2 General Activity Schedule for Conducting a Long-term Sediment Toxicity Test with *Chironomus tentans*

Day	Activity
Pre-Test	
-4	Start reproduction flask with cultured adults (1:3 male:female ratio). For example for 15 to 25 egg cases, 10 males and 30 females are typically collected. Egg cases typically range from 600 to 1500 eggs/case.
-3	Collect egg cases (a minimum of 6 to 8) and incubate at 23°C.
-2	Check egg cases for viability and development.
-1	1. Check egg cases for hatch and development.
	2. Add 100 mL of homogenized test sediment to each replicate beaker and place in corresponding treatment holding tank. After sediment has settled for at least 1 h, add 1.5 mL Tetrafin slurry (4g/L solution) to each beaker. Overlying water renewal begins at this time.
Sediment Test	
0	1. Transfer all egg cases to a crystallizing dish containing control water. Discard larvae that have already left the egg cases in the incubation dishes. Add 1.5 mL food to each test beaker with sediment before the larvae are added. Add 12 larvae to each replicate beaker (beakers are chosen by random block assignment). Let beakers sit (outside the test system) for 1 h following addition of the larvae. After this period, gently immerse all beakers into their respective treatment holding tanks.
	2. Measure temperature, pH, hardness, alkalinity, dissolved oxygen, conductivity and ammonia at start of test.
1-End	On a daily basis, add 1.5 mL food to each beaker. Measure temperature daily. Measure the pH and dissolved oxygen three times a week during the test. Measure conductivity weekly. If the DO has declined more than 1 mg/L since previous reading, increase frequency of DO measurements and aerate if DO continues to be less than 2.5 mg/L. Measure hardness, alkalinity, conductivity, ammonia, temperature, pH, and dissolved oxygen at the end of the test.
6	For auxiliary male production, start reproduction flask with culture adults (e.g., 10 males and 30 females; 1:3 male to female ratio).
7-10	Follow set-up schedule for auxiliary male beakers (4 replicates/treatment) described above for Day -3 to Day 0.
19	In preparation for weight determinations, ash weigh pans at 550°C for 2 h. Note that the weigh pans should be ashed before use to eliminate weighing errors due to the pan oxidizing during ashing of samples.
20	1. Randomly select four replicates from each treatment and sieve the sediment to recover larvae for growth and survival determinations. Pool all living larvae per replicate and dry the sample to a constant weight (e.g., 60°C for 24 h).
	2. Install emergence traps on each of the remaining reproductive replicate beakers.
	3. Measure temperature, pH, hardness, alkalinity, dissolved oxygen, conductivity and ammonia.
21	The sample with dried larvae is brought to room temperature in a dessicator and weighed to the nearest 0.01 mg . The dried larvae in the pan are then ashed at 550ºC for 2 h. The pan with the ashed larvae is then reweighed and the tissue mass of the larvae determined as the difference between the weight of the dried larvae plus pan and the weight of the ashed larvae plus pan.
Chronic Measurements	
23-End	On a daily basis, record emergence of males and females, pupal, and adult mortality, and time to death for previously collected adults. Each day, transfer adults from each replicate to a corresponding reproduction/oviposition (R/O) chamber. Transfer each primary egg case from the R/O chamber to a corresponding petri dish to monitor incubation and hatch. Record each egg case oviposited, number of eggs produced (using either the ring or direct count methods), and number of hatched eggs. If it is difficult to estimate the number of eggs in an egg case, use a direct count to determine the number of eggs; however the hatchability data will not be obtained for this egg case.
28	Place emergence traps on auxiliary male replicate beakers.
33-End	Transfer males emerging from the auxiliary male replicates to individual inverted petri dishes. The auxiliary males are used for mating with females from corresponding treatments from which most of the males had already emerged or in which no males emerged.
40-End	After 7 d of no recorded emergence in a given treatment, end the treatment by sieving the sediment to recover larvae, pupae, or pupal exuviae. When no emergence occurs in a test treatment, that treatment can be ended once emergence in the control sediment has ended using the 7-d criterion.

replicates for determination of emergence and reproduction. It is typical for males to begin emerging 4 to 7 d before females. Therefore, additional males, referred to as auxiliary males, need to be available during the prime female emergence period for each respective chamber/sediment. To provide these males, 4 additional replicates are stocked with 12, <24-h-old larvae 10 d following initiation of the test. Midges in each test chamber are fed 1.5 mL of a 4-g/L Tetrafin® suspension daily. Endpoints monitored include 20-d survival and weight, emergence, time to death (adults), reproduction, and egg hatchability.

Table 15.3 Test Acceptability Requirements for a Long-term Sediment Toxicity Test with *Chironomus tentans*

A. It is recommended for conducting a long-term test with *C. tentans* that the following performance criteria be met:

1. Tests must be started with less than 1-d- (<24-h) old larvae. Starting a test with substantially older organisms may compromise the emergence and reproductive endpoint.
2. Average survival of *C. tentans* in the control sediment should be greater than or equal to 70% on Day 20 and greater than 65% at the end of the test.
3. Average size of *C. tentans* in the control sediment at 20 d must be at least 0.6 mg/surviving organism as dry weight or 0.48 mg/surviving organism as AFDW. Emergence should be greater than or equal to 50%. Experience has shown that pupae survival is typically >83% and adult survival is >96%. Time to death after emergence is <6.5 d for males and <5.1 d for females. The mean number of eggs/egg case should be greater than or equal to 800 and the percent hatch should be greater than or equal to 80%. See Sections 15.1.3 and 17.6 for a summary of performance in round-robin testing.
4. Hardness, alkalinity, and ammonia in the overlying water typically should not vary by more than 50% during the test, and dissolved oxygen should be maintained above 2.5 mg/L in the overlying water.

B. Performance-based criteria for culturing *C. tentans* include the following:

1. It may be desirable for laboratories to periodically perform 96-h water-only reference-toxicity tests to assess the sensitivity of culture organisms (Section 9.16.2). Data from these reference-toxicity tests could be used to assess genetic strain or life-stage sensitivity of test organisms to select chemicals.
2. Laboratories should keep a record of time to first emergence for each culture and record this information using control charts. Records should also be kept on the frequency of restarting cultures.
3. Laboratories should record the following water-quality characteristics of the cultures at least quarterly: pH, hardness, alkalinity, and ammonia. Dissolved oxygen in the cultures should be measured weekly. Temperature of the cultures should be recorded daily. If static cultures are used, it may be desirable to measure water quality more frequently.
4. Laboratories should characterize and monitor background contamination and nutrient quality of food if problems are observed in culturing or testing organisms.
5. Physiological measurements such as lipid content might provide useful information regarding the health of the cultures.

C. Additional requirements:

1. All organisms in a test must be from the same source.
2. Storage of sediments collected from the field should follow guidance outlined in Section 8.2.
3. All test chambers (and compartments) should be identical and should contain the same amount of sediment and overlying water.
4. Negative-control sediment and appropriate solvent controls must be included in a test. The concentration of solvent used must not adversely affect test organisms.
5. Test organisms must be cultured and tested at 23°C (±1°C).
6. The daily mean test temperature must be within ±1°C of 23°C. The instantaneous temperature must always be within ±3°C of 23°C.
7. Natural physico-chemical characteristics of test sediment collected from the field should be within the tolerance limits of the test organisms.

15.3 General Procedures

15.3.1 Collection of Egg Cases

15.3.1.1 Egg cases are obtained from adult midges held in a sex ratio of 1:3 male:female. Ten males and 30 females will produce between 15 to 25 egg cases. Adults should be collected four days before starting a test (Appendix C, Figure C.3). The day after collection of adults, 6 to 8 of the larger “C” shaped egg cases are transferred to a petri dish with culture water and incubated at 23°C (Appendix C, Figure C.2). Hatching typically begins around 48 h and larvae typically leave the egg case 24 h after the first hatch. The number of eggs in each egg case will vary, but typically ranges from 600 to 1500 eggs. It should be noted that mating may have occurred in culture tanks before males and females are placed into flasks for collecting eggs.

15.3.2 Hatching of Eggs

15.3.2.1 Hatching of eggs should be complete by about 72 h. Hatched larvae remain with the egg case for about 24 h and appear to use the gelatinous component of the egg case as an initial source of food (Sadler, 1935; Ball and Baker, 1995). After the first 24-h period with larvae hatched, transfer the egg cases from the incubation petri dish to another dish with clean test water. Larvae having already left the egg case in the incubation petri dish are discarded since their precise age and time away from the gelatinous food source is unknown. The action of transferring the egg case stimulates the remaining larvae to leave the egg case within a few hours. These are the larvae that are used to start the test.

Table 15.4 Endpoints for a Long-term Sediment Toxicity Test with *Chironomus tentans*

Lethal	Sublethal		
Survival	Growth	Emergence	Reproduction
Larvae (20 d)	Larvae	Total/Percent	Sex Ratio
Larvae (End)		Cumulative (Rate)	Time to Oviposition
Pupae		Time to First	Mean Eggs/Female
Adults		Time to Death	Egg Cases/Treatment
			Egg Hatchability

15.3.3 Sediment into Test Chambers

15.3.3.1 The day before the sediment test is started (Day -1) each sediment should be thoroughly homogenized and added to the test chambers (Section 8.3.1). Sediment should be visually inspected to judge the extent of homogeneity. Excess water on the surface of the sediment can indicate separation of solid and liquid components. If a quantitative measure of homogeneity is required, replicate subsamples should be taken from the sediment batch and analyzed for TOC, chemical concentrations, and particle size.

15.3.3.2 Each test chamber should contain the same amount of sediment, determined either by volume or by weight. Overlying water is added to the chambers in a manner that minimizes suspension of sediment. This can be accomplished by gently pouring water along the sides of the chambers or by pouring water onto a baffle (e.g., a circular piece of Teflon with a handle attached) placed above the sediment to dissipate the force of the water. Renewal of overlying water is started on Day -1. A test begins when the organisms are added to the test chambers (Day 0).

15.3.4 Renewal of Overlying Water

15.3.4.1 Renewal of overlying water is required during a test. Two volume additions of overlying water (continuous or intermittent) should be delivered to each test chamber daily. At any particular time during the test, flow rates through any two test chambers should not differ by more than 10%. Hardness, alkalinity and ammonia concentrations in the water above the sediment, within a treatment, typically should not vary by more than 50% during the test. Mount and Brungs (1967) diluters have been modified for sediment testing, and other automated water-delivery systems have also been used (Maki, 1977; Ingersoll and Nelson, 1990; Benoit et al., 1993; Zumwalt et al., 1994; Brunson et al., 1998; Wall et al., 1998; Leppanen and Maier, 1998). Each water-delivery system should be calibrated before a test is started to verify that the system is functioning properly. Renewal of overlying water is started on Day -1 before the addition of test organisms on Day 0. Appendix A describes water-renewal systems that can be used for conducting sediment tests.

15.3.4.2 In water-renewal tests with one to four volume additions of overlying water/d, water-quality characteristics generally remain similar to the inflowing water (Ingersoll and Nelson, 1990; Ankley et al., 1993); however, in static tests, water quality may change profoundly during the exposure (Shuba et al., 1978). For example, in static whole-sediment tests, the alkalinity, hardness, and conductivity of overlying water more than doubled in several treatments during a four-week exposure (Ingersoll and Nelson, 1990). Additionally, concentrations of metabolic products (e.g., ammonia) may also increase during static exposures, and these compounds can either be directly toxic to the test organisms or may contribute to the toxicity of the contaminants in the sediment. Furthermore, changes in water-quality characteristics such as hardness may influence the toxicity of many inorganic (Gauss et al., 1985) and organic (Mayer and Ellersieck, 1986) contaminants. Although contaminant concentrations are reduced in the overlying water in water-renewal tests, organisms in direct contact with sediment generally receive a substantial proportion of a contaminant dose directly from either the whole sediment or from the interstitial water.

15.3.5 Acclimation

15.3.5.1 Test organisms must be cultured and tested at 23°C. Ideally, test organisms should be cultured in the same water that will be used in testing. However, acclimation of test organisms to the test water is not required.

15.3.5.2 Culturing of organisms and toxicity assessment are typically conducted at 23°C. However, occasionally there is a need to perform evaluations at temperatures different than that recommended. Under these circumstances, it may be necessary to acclimate organisms to the desired test temperature to prevent thermal shock when moving immediately from the culture temperature to the test temperature (ASTM, 1999a). Acclimation can be achieved by exposing organisms to a gradual decline in temperature; however, the rate of decline should be relatively slow to prevent thermal shock. A decline in temperature of 1°C every 1 to 2 h has been used successfully in some studies (P.K. Sibley, University of Guelph, Guelph, Ontario, personal communication; APHA, 1989). Testing at temperatures other than 23°C needs to be preceded by

studies to determine expected performance under alternate conditions.

15.3.6 Placing Organisms in Test Chambers

15.3.6.1 Test organisms should be handled as little as possible. To start the test, larvae are collected with a Pasteur pipet from the bottom of the incubation dish with the aid of a dissecting microscope. Test organisms are pipetted directly into overlying water and care should be exercised to release them under the surface of the water. Transferring the larvae to exposure chambers within 4 h of emerging from the egg case reportedly improves survival (Benoit et al., 1997). Laboratory personnel should practice transferring first-instar midge larvae before tests with sediment are conducted.

15.3.7 Feeding

15.3.7.1 Each beaker receives a daily addition of 1.5 mL of Tetrafin® (4 mg/mL dry solids). Without addition of food, the test organisms may starve during exposures. However, the addition of the food may alter the availability of the contaminants in the sediment (Wiederholm et al., 1987; Harkey et al., 1994). Furthermore, if too much food is added to the test chamber, or if the mortality of test organisms is high, fungal or bacterial growth may develop on the sediment surface. Therefore, the amount of food added to the test chambers is kept to a minimum.

15.3.7.1 Suspensions of food should be thoroughly mixed before aliquots are taken. If excess food collects on the sediment, a fungal or bacterial growth may develop on the sediment surface, in which case feeding should be suspended for one or more days. A drop in dissolved oxygen below 2.5 mg/L during a test may indicate that the food added is not being consumed. Feeding should be suspended for the amount of time necessary to increase the dissolved oxygen concentration (ASTM, 1999a). If feeding is suspended in one treatment, it should be suspended in all treatments. Detailed records of feeding rates and the appearance of the sediment surface should be made daily.

15.3.8 Monitoring a Test

15.3.8.1 All chambers should be checked daily and observations made to assess test organism behavior such as sediment avoidance. However, monitoring effects on burrowing activity of test organisms may be difficult because the test organisms are often not visible during the exposure. The operation of the exposure system should be monitored daily.

15.3.8.2 Measurement of Overlying Water-quality Characteristics

15.3.8.2.1 Conductivity, hardness, alkalinity, and ammonia should be measured in all treatments at the beginning of the test, on Day 20, and at the end of the test. Dissolved oxygen (DO) and pH measurements should be taken at the beginning of a test and at least three times a week until the end of the test. Conductivity should be measured weekly. Overlying water should be sampled just before water renewal from about 1 to 2 cm above the sediment surface using a pipet. It may be necessary to composite water samples from individual replicates. The pipet should be checked to make sure no organisms are removed during sampling of overlying water. Water quality should be measured on each batch of water prepared for the test.

15.3.8.2.2 Routine chemistries on Day 0 should be taken before organisms are placed in the test beakers. Dissolved oxygen and pH can be measured directly in the overlying water with a probe. However, for DO it is important to allow the probe time to equilibrate in the overlying water in an effort to accurately measure concentrations of DO. If a probe is used for measurements in overlying water, it should be inspected between samples to make sure that organisms are not attached and should be rinsed between samples to minimize cross contamination.

15.3.8.2.3 Water-only exposures evaluating the tolerance of *C. tentans* larva to depressed DO have indicated that significant reductions in weight occurred after 10-d exposure to 1.1 mg/L DO, but not at 1.5 mg/L (V. Mattson, USEPA, Duluth, MN, personal communication). This finding concurs with the observations during method development at the USEPA laboratory in Duluth that excursions of DO as low as 1.5 mg/L did not seem to have an effect on midge survival and development (P.K. Sibley, University of Guelph, Guelph, Ontario, personal communication). Based on these findings, periodic depressions of DO below 2.5 mg/L (but not below 1.5 mg/L) are not likely to adversely affect test results, and thus should not be a reason to discard test data. Nonetheless, tests should be managed toward a goal of DO >2.5 mg/L to insure satisfactory performance. If the DO level of the water falls below 2.5 mg/L for any one treatment, aeration is encouraged and should be done in all replicates for the duration of the test (i.e., about 1 bubble/second in the overlying water). Occasional brushing of screens on outside of beakers will help maintain the exchange of water during renewals.

15.3.8.2.4 Temperature should be measured at least daily in at least one test chamber from each treatment. The temperature of the water bath or the exposure chamber should be continuously monitored. The daily mean test temperature must be within ±1°C of 23°C. The instantaneous temperature must always be within ±3°C of 23°C.

15.3.8.3 Monitoring Survival and Growth

15.3.8.3.1 At 20 d, 4 of the initial 12 replicates are selected for use in growth and survival measurements. Using a #40 sieve (425-µm mesh) to remove larvae from sediment, collect the *C. tentans* and record data on record sheet (Appendix D). Any immobile organisms isolated from the sediment surface or from sieved material should be considered dead. Often *C. tentans* larvae tend to lose their coloration within 15 to 20 min of death and may become rigidly elongate. Surviving larvae are kept separated by replicate for weight measurements; if pupae are

recovered (<1% occurrence at recommended testing conditions), these organisms are included in survival data but not included in the growth data. A consistent amount of time should be taken to examine sieved material for recovery of test organisms (e.g., 5 min/replicate).

15.3.8.3.2 The 10-d method for *C. tentans* in the first edition of this manual (USEPA, 1994a), as well as most previous research, has used dry weight as a measure of growth. However, Sibley et al. (1997b) found that the grain size of sediments influences the amount of sediment that *C. tentans* larvae ingest and retain in their gut. As a result, in finer-grain sediments, a substantial portion of the measured dry weight may be comprised of sediment rather than tissue. While this may not represent a strong bias in tests with identical grain size distributions in all treatments, most field assessments are likely to have varying grain size among sites. This will likely create differences in dry weight among treatments that are not reflective of true somatic growth. For this reason, weight of midges should be measured as ash-free dry weight (AFDW) instead of dry weight. AFDW will more directly reflect actual differences in tissue weight by reducing the influence of sediment in the gut. If test organisms are to be used for an evaluation of bioaccumulation, it is not advisable to dry the sample before conducting the residue analysis. If conversion from wet weight to dry weight is necessary, aliquots of organisms can be weighed to establish wet to dry weight conversion factors. A consistent procedure should be used to remove the excess water from the organisms before measuring wet weight.

15.3.8.3.3 The AFDW of midges should be determined for the growth endpoint. All living larvae per replicate are combined and dried to a constant weight (e.g., 60°C for 24 h). Note that the weigh boats should be ashed before use to eliminate weighing errors due to the pan oxidizing during ashing. The sample is brought to room temperature in a desiccator and weighed to the nearest 0.01 mg to obtain mean weights per surviving organism per replicate. The dried larvae in the pan are then ashed at 550°C for 2 h. The pan with the ashed larvae is then reweighed and the tissue mass of the larvae is determined as the difference between the weight of the dried larvae plus pan and the weight of the ashed larvae plus pan. For rare instances in which preservation is required, an 8% sugar formalin solution can be used to preserve samples (USEPA, 1994a), but the effects of preservation on the weight and lengths of the midges have not been sufficiently studied. The sugar formalin solution is prepared by adding 120 g of sucrose to 80 mL of formalin which is then brought to a volume of 1 L using deionized water. This stock solution is mixed with an equal volume of deionized water when used to preserve organisms. NoTox® (Earth Safe Industries, Belle Mead, NJ) can be used as a substitute for formalin (Unger et al., 1993).

15.3.8.4 Monitoring Emergence

15.3.8.4.1 Emergence traps are placed on the reproductive replicates on Day 20 (emergence traps for the auxiliary beakers are added at the corresponding 20-d time interval for those replicates; Appendix C, Figures C.1 and C.4). At 23 °C, emergence in control sediments typically begins on or about Day 23 and continues for about 2 weeks. However, in contaminated sediments, the emergence period may be extended by several weeks.

15.3.8.4.2 Two categories are recorded for emergence: complete emergence and partial emergence. Complete emergence occurs when an organism has shed the pupal exuviae completely and escapes the surface tension of the water. If complete emergence has occurred but the adult has not escaped the surface tension of the water, the adult will die within 24 h. Therefore, 24 h should elapse before this death is recorded. Partial emergence occurs when an adult has only partially shed the pupal exuviae. These adults will also die, an event which can be recorded after 24 h. Pupae at the sediment surface or the air-water interface may emerge successfully during the 24-h period. However, cannibalism of sediment bound pupae by larvae may also occur. Data are recorded on data sheets provided as shown in example data sheet (Appendix D).

15.3.8.4.3 Between Day 23 and the end of the test, emergence of males and females, pupal and adult mortality, and time to death for adults is recorded daily for the reproductive replicates. On Day 30 (20-d-old organisms), emergence traps are placed on the auxiliary beakers to collect the additional males for use with females emerging from the reproduction replicates (Table 15.2; Appendix C, Figures C.1 and C.4). Data are recorded on data sheets provided as shown in the example data sheet (Appendix D).

15.3.8.5 Collecting Adults for Reproduction

15.3.8.5.1 Adults are collected daily from individual traps using the aspirator and collector dish (Appendix C, Figure C.2). With the collector dish nearby, the emergence trap is quickly moved from the beaker onto the dish. With the syringe plunger fully drawn, the glass collector tube is inserted through the screened access hole of the collector dish and the adults gently aspirated into the syringe barrel. Aspirated adults can easily be seen through the translucent plastic of the syringe. The detachable portion of the aspirator unit is then replaced with a reproduction/oviposit (R/O) chamber. This exchange can be facilitated by placing the thumb of the hand holding the syringe over the barrel entry port until the R/O chamber is in place. With the R/O chamber in place, and the plunger on a solid surface, the barrel of the syringe is pushed gently downward which forces the adults to move up into the R/O unit. Adults remaining on the transfer apparatus may be prodded into the R/O chamber by gently tapping the syringe. The transfer process is completed by quickly moving the R/O chamber to a petri dish containing clean water. At all times during

the transfer process, it is important to ensure that the adults are stationary to minimize the possibility of escape.

15.3.8.5.2 At about Day 33 to the end of the test, the auxiliary males may be needed to support reproduction in females. Males that emerge from the auxiliary male replicates are transferred to individual inverted petri dishes (60 x 15 mm dishes without water and with air holes drilled in top of the dish; see Appendix C for a listing of equipment.) Each male may be used for mating with females from corresponding treatments for up to 5 d. Males may be used for breeding with more than one new emergent female. Males from a different replicate within the same sediment treatment may be paired with females of replicates where no males have emerged. Data can be recorded on data sheets provided in Appendix D.

15.3.8.6 Monitoring Reproduction

15.3.8.6.1 Each R/O unit is checked daily for dead adults and egg cases. Dead organisms are removed. In situations where many adults are contained within an R/O chamber, it may be necessary to assume that a dead adult is the oldest male or female in that replicate for the purpose of recording time to death. To remove dead adults and egg cases from the R/O chamber, one side of the chamber is carefully lifted just enough to permit the insertion of a transfer pipet or tweezers.

15.3.8.6.2 For each emerged female, at least one male, obtained from the corresponding reproductive replicate, from another replicate of that treatment, or from the auxiliary male beakers, is transferred into the R/O unit using an aspirator. Females generally remain sexually receptive up to 3 d if they have not already mated. Benoit et al. (1997) have shown that over 90% of females will oviposit within 1 d of fertilization; however, a few will require as long as 72 h to oviposit. A female will lay a single primary egg case, usually in the early morning (Sadler, 1935). A second, generally smaller egg case may be laid; however these second egg cases are prone to fungus and the viability of embryos is typically poor. These second egg cases do not need to be counted, or recorded, and the numbers of eggs are not included in the egg counts because eggs in second egg cases typically have lower viability.

15.3.8.7 Counting Eggs, Egg Case Incubation, and Hatch Determination

15.3.8.7.1 Primary egg cases from the R/O chamber are transferred to a separate and corresponding petri dish (60 x 15 mm with about 15 mL of water) to monitor incubation and hatch. The number of eggs should be estimated in each egg case by using a "ring method" as follows: (1) for each egg case, the mean number of eggs in five rings is determined; (2) these rings should be selected at about equal distances along the length of the egg case; (3) the number of eggs/ring multiplied by the number of number of rings in the egg case will provide an estimate of the total number of eggs. This can be done in about 5 min or less for each egg case. Accuracy of estimating versus a direct count method is very close, roughly 95% (Benoit et al., 1997). The ring method is best suited to the “C” shaped egg cases.

15.3.8.7.2 When the integrity of an egg case precludes estimation by the ring method (egg case is convoluted or distorted), the eggs should be counted directly. Each egg case is placed into a 5-cm glass culture tube containing about 2 mL of 2 N sulfuric acid (H_2SO_4) and left overnight. The acid dissolves the gelatinous matrix surrounding the eggs but does not affect the structural integrity of the eggs themselves. After digestion, the eggs are collected with a Pasteur pipet and spread across a microscope slide for counting under a dissecting microscope. Counting can be simplified by drawing a grid on the underside of the slide. The direct count method requires a minimum of 10 min to complete and does not permit determination of hatching success.

15.3.8.7.3 Following estimated egg counts, each egg case is transferred to a 60- x 15-mm plastic petri dish containing 15 mL overlying water and incubated at 23°C until hatching is complete. Although the time required to initiate hatching at this temperature is about 2 d, the period of time required to bring about complete hatch may be as long as 6 d. Therefore, hatching success is determined after 6 d of incubation. Hatching success is determined by subtracting the number of unhatched eggs remaining after the 6 d period from the number of eggs originally estimated for that egg case. Unhatched eggs either remain in the gelatinous egg case or are distributed on the bottom of the petri dish.

15.3.8.7.4 Depending on the objectives of the study, reproductive output in *C. tentans* may be expressed as: (1) number of eggs/female or (2) number of offspring/female. The former approach estimates reproductive output (fecundity) in terms of the number of eggs deposited by a female (secondary egg cases are not included) and does not take into account survival of hatched eggs. This approach has been shown to adequately discriminate contaminant (Sibley et al., 1996) and noncontaminant (Sibley et al., 1997a) stressors. Since this approach does not require monitoring egg masses for hatchability, the time and labor involved in conducting the life-cycle test is reduced. However, studies that require estimates of demographic parameters, or include population modeling, will need to determine the number of viable offspring per female (Sibley et al., 1997a). This will require determination of larval hatch (see Section 15.3.8.7.3). Although larval hatch is listed as a potential endpoint by itself in this manual (Table 15.4), the sensitivity of this endpoint has not been fully assessed.

15.3.9 Ending a Test

15.3.9.1 The point at which the life-cycle test is ended depends upon the sediments being evaluated. In clean sediments, the test typically requires 40 to 50 d from initial setup to completion. However, test duration will increase in the presence of environmental stressors which

act to reduce growth and delay emergence (Sibley et al., 1997a). Where a strong gradient of sediment contamination exists, emergence patterns between treatments will likely become asynchronous, in which case each treatment needs to be ended separately. For this reason, emergence is used as a guide to decide when to end a test.

15.3.9.2 For treatments in which emergence has occurred, the treatment (not the entire test) is ended when no further emergence is recorded over a period of 7 d (the 7-d criterion). At this time, all beakers of the treatment are sieved through a #40-mesh screen (425 µm) to recover remaining larvae, pupae, or pupal castes. When no emergence is recorded in a treatment at any time during the test, that treatment can be ended once emergence in the control sediment has ended using the 7-d criterion.

15.4 Interpretation of Results

15.4.1 Data Analysis

15.4.1.1 Endpoints measured in the *C. tentans* test include survival, growth, emergence and reproduction. Section 16 describes general information regarding statistical analysis of these data, including both point estimates (i.e., LC50s) and hypothesis testing (i.e., ANOVA). The following sections describe species-specific information that is useful in helping to interpret the results of long-term sediment toxicity tests with *C. tentans*.

15.4.2 Age Sensitivity

15.4.2.1 Midges are perceived to be relatively insensitive organisms in toxicity assessments (Ingersoll, 1995). This conclusion is based on the practice of measuring survival of fourth-instar larvae in short-term water-only exposures, a procedure that may underestimate the sensitivity of midges to toxicants. The first and second instars of chironomids are more sensitive to contaminants than the third or fourth instars. For example, first-instar *C. tentans* larvae were 6 to 27 times more sensitive than fourth-instar larvae to acute copper exposure (Nebeker et al., 1984b; Gauss et al., 1985; Figure 12.1) and first-instar *C. riparius* larvae were 127 times more sensitive than second-instar larvae to acute cadmium exposure (Williams et al., 1986b; Figure 12.1). In long-term tests with first-instar larvae, midges were often as sensitive as daphnids to inorganic and organic compounds (Ingersoll et al., 1990). Sediment tests should be started with uniform age and size midges because of the dramatic differences in sensitivity of midges by age.

15.4.3 Physical Characteristics of Sediment

15.4.3.1 Grain Size

15.4.3.1.1 Larvae of *C. tentans* appear to be tolerant of a wide range of particle size conditions in substrates. Several studies have shown that survival is not affected by particle size in natural sediments, sand substrates, or formulated sediments in both 10-d and long-term exposures (Ankley et al., 1994; Suedel and Rodgers, 1994; Sibley et al., 1997b, 1998). Ankley et al. (1994a) found that growth of *C. tentans* larvae was weakly correlated with sediment grain size composition, but not organic carbon, in 10-d tests using 50 natural sediments from the Great Lakes. However, Sibley et al. (1997b) found that the correlation between grain size and larval growth disappeared after accounting for inorganic material contained within larval guts and concluded that growth of *C. tentans* was not related to grain size composition in either natural sediments or sand substrates. Avoiding confounding influences of gut contents on weight is the impetus for recommending ash-free dry weight (instead of dry weight) as the index of growth in the 10-day and long-term *C. tentans* tests. Failing to do so could lead to erroneous conclusions regarding the toxicity of the test sediment (Sibley et al., 1997b). Procedures for correcting for gut contents are described in Section 15.3.8.3. Emergence, reproduction (mean eggs/female), and hatch success were also not affected by the particle size composition of substrates in long-term tests with *C. tentans* (Sibley et al., 1998).

15.4.3.2 Organic Matter

15.4.3.2.1 Based on 10-d tests, the content of organic matter in sediments does not appear to affect survival of *C. tentans* larvae in natural and formulated sediments, but may be important with respect to larval growth. Ankley et al. (1994a) found no relationship between sediment organic content and survival or growth in 10-d bioassays with *C. tentans* in natural sediments. Suedel and Rodgers (1994) observed reduced survival in 10-d tests with a formulated sediment when organic matter was <0.91%; however, supplemental food was not supplied in this study, which may influence these results relative to the 10-d test procedures described in this manual. Lacey et al. (1999) found that survival of *C. tentans* larvae was generally not affected in 10-d tests by either the quality or quantity of synthetic (alpha-cellulose) or naturally derived (peat, maple leaves) organic material spiked into a formulated sediment, although a slight reduction in survival below the acceptability criterion (70%) was observed in a natural sediment diluted with formulated sediment at an organic matter content of 6%. In terms of larval growth, Lacey et al. (1999) did not observe any systematic relationship between the level of organic material (e.g., food quantity) and larval growth for each carbon source. Although a significant reduction in growth was observed at the highest concentration (10%) of the leaf treatment in the food quantity study, significantly higher larval growth was observed in this treatment when the different carbon sources were compared at about equal concentrations (effect of food quality). In the latter study, the following gradient of larval growth was established in relation to the source of organic carbon: peat < natural sediment < alpha-cellulose < leaves. Since all of the treatments received a supplemental source of food, these data suggest that both the quality and quantity of organic carbon in natural and formulated sediments may represent an important confounding factor for the growth endpoint in tests with *C. tentans* (Lacey et al., 1999). However, it is

important to note that these data are based on 10-d tests; the applicability of these data to long-term testing has not been evaluated.

15.4.4 Isolating Organisms at the End of a Test

15.4.4.1 Quantitative recovery of larvae at the end of a sediment test should not be a problem. The larvae are red and typically greater than 5 mm long and are readily retained on the #40-mesh sieve.

15.4.5 Influence of Indigenous Organisms

15.4.5.1 The influence of indigenous organisms on the response of *C. tentans* in sediment tests has not been reported. Survival of a closely related species, *C. riparius* was not reduced in the presence of oligochaetes in sediment samples (Reynoldson et al., 1994). However, growth of *C. riparius* was reduced when high numbers of oligochaetes were placed in a sample. Therefore, it is important to determine the number and biomass of indigenous organisms in field-collected sediment in order to better interpret growth data (Reynoldson et al., 1994; DeFoe and Ankley, 1998). Furthermore, the presence of predators may also influence the response of test organisms in sediment (Ingersoll and Nelson, 1990).

15.4.6 Relationship Between Endpoints

15.4.6.1 Relationship Between Growth and Emergence Endpoints

15.4.6.1.1 An important stage in the life cycle of *C. tentans* is the emergence of adults from pupal forms. Emergence has been used in many studies as an indicator of contaminant stress (Wentsel et al., 1978; Pascoe et al., 1989; Sibley et al., 1996). The use of emergence as an endpoint in this context is based upon the understanding that larval growth and emergence are intimately related such that environmental factors that affect larval development may also affect emergence success. Implicit in the relationship between growth and emergence is the notion of a weight threshold that needs to be attained by larvae in order for emergence to take place (Hilsenhoff,1966; Liber et al., 1996; Sibley et al., 1997a). For example, based on evaluations conducted in clean control sediment, Liber et al. (1996) and Sibley et al. (1997a) showed that a minimum tissue mass threshold of approximately 0.6 mg dry weight or 0.48 mg ash-free dry weight was required before pupation and emergence could take place (Figure 15.1). Further, Sibley et al. (1997a) found that maximum emergence (e.g., >60%) in this sediment occurred only after larvae had attained a tissue mass of about 0.8 mg dry weight. This value corresponds closely to that suggested by Ankley et al. (1994a) as an acceptability criterion for growth in control sediments in 10-d tests with *C. tentans*.

15.4.6.2 Relationship Between Growth and Reproduction Endpoints

15.4.6.2.1 Natural or anthropogenic stressors that affect growth of invertebrates may also affect reproduction, because of a minimum threshold body mass needed for reproduction (Rees and Crawley, 1989; Ernsting et al., 1993; Moore and Dillon, 1993; Sibley et al., 1996,1997a). Sibley et al. (1996,1997a) reported a significant relationship between growth (dry weight) of larval *C. tentans* and reproductive output (mean number of eggs) of adults in relation to both food and contaminant (zinc) stressors (Figure 15.2). The form that this relationship may take depends upon the range of stress to which the larvae are exposed and may be linear or sigmoidal. The latter relationship is typically characterized by an upper maximum determined by competitive factors (i.e., food and space availability) and a lower minimum determined primarily by emergence thresholds (See Section 15.4.6.1; Sibley et al., 1997a).

15.4.6.2.2 Embryo viability (percent hatch of eggs) has been shown to evaluate the toxicity for waterborne chemicals (Williams et al.,1986b; Pascoe et al.,1989). However, percent hatch has not been used extensively as an endpoint to assess toxicity in contaminated sediments. Sibley et al. (1996) found that the viability of embryos was not affected at any of the zinc treatments for which egg masses were produced; >87% of all eggs eventually hatched. Additional information regarding the measurement of embryo viability in round-robin testing is presented in Section 17.6.

15.4.6.2.3 In contrast to *H. azteca* (Section 14.4), length is not commonly utilized as a growth endpoint in *C. tentans*. However, length may represent a useful alternative to weight. For example, recent studies (P.K. Sibley, University of Guelph, Guelph, Ontario, unpublished data) found a significant relationship (r^2=0.99; $p < 0.001$) between ash-free dry weight and length in larvae of *C. tentans* reared in clean control sediment (Figure 15.3). This suggests that either weight or length could be used to assess growth in *C. tentans*. However, the relationship between length and emergence or reproductive endpoints has not been evaluated.

15.4.6.3 Relationship Between Growth and Population Endpoints

15.4.6.3.1 Few studies have attempted to quantitatively define the relationship between larval growth and population-level processes. However, an accurate understanding of the ecological relevance of growth as an endpoint in sediment toxicity tests can only be achieved in terms of its effect, if any, on population-level processes. Sibley et al. (1997a) found a significant relationship between larval growth and the intrinsic rate of population increase in *C. tentans* in relation to a food stressor (Figure 15.4). When applied in a theoretical population model, it was further demonstrated that changes in larval growth resulting from the stressor gradient were significantly correlated to the predicted number of offspring recruited to subsequent generations.

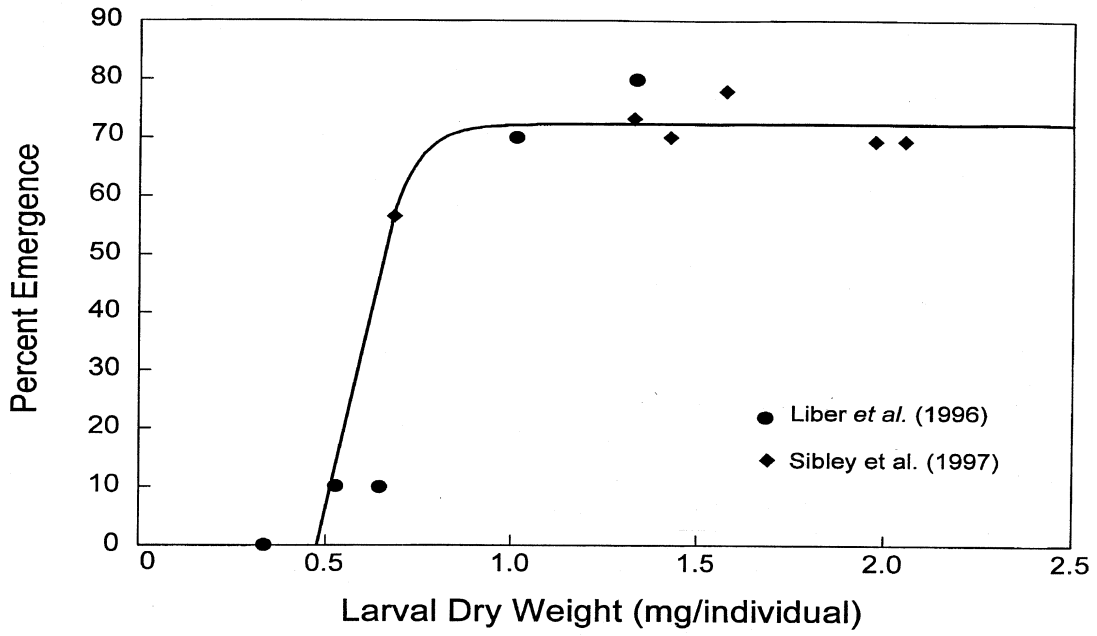


Figure 15.1 Relationship between weight and emergence of *Chironomus tentans.*

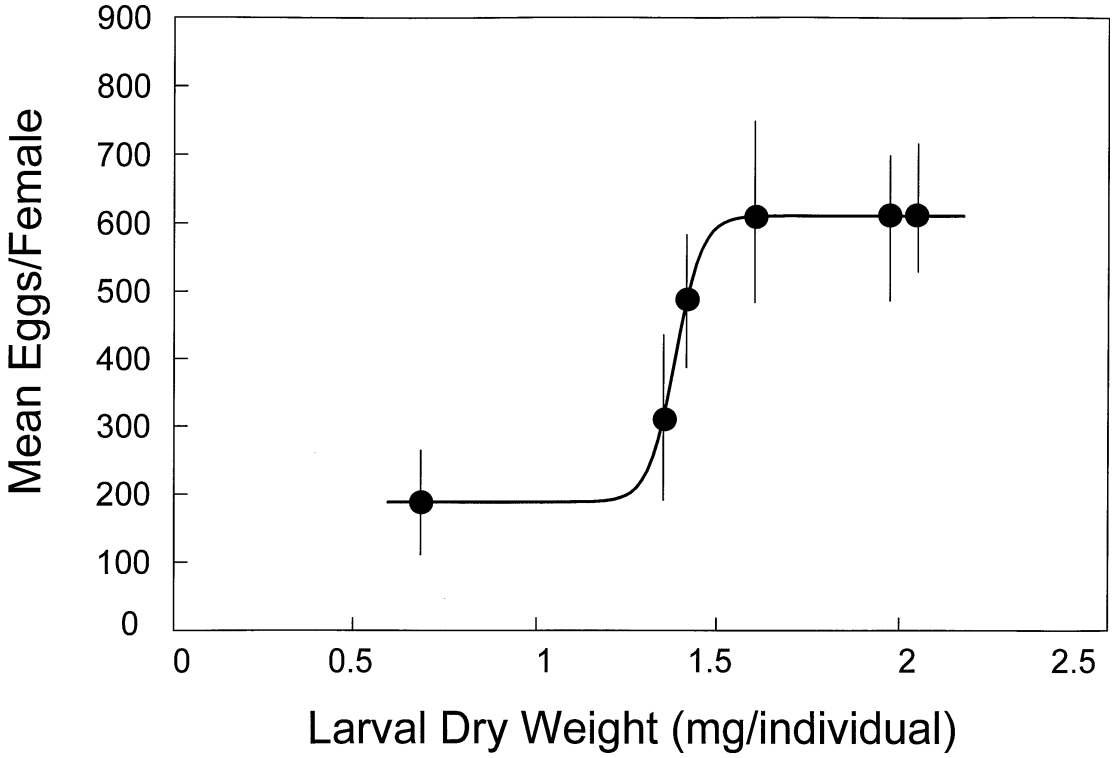


Figure 15.2 Relationship between weight and reproduction of *Chironomus tentans.*

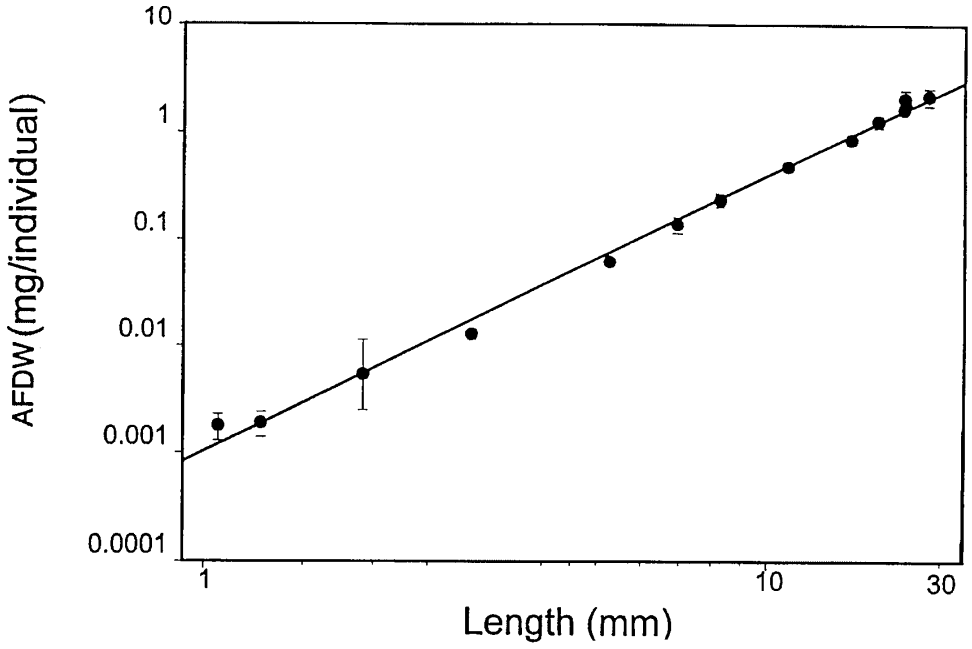


Figure 15.3 Relationship between ash-free dry weight (AFDW) and length of *Chironomus tentans.*

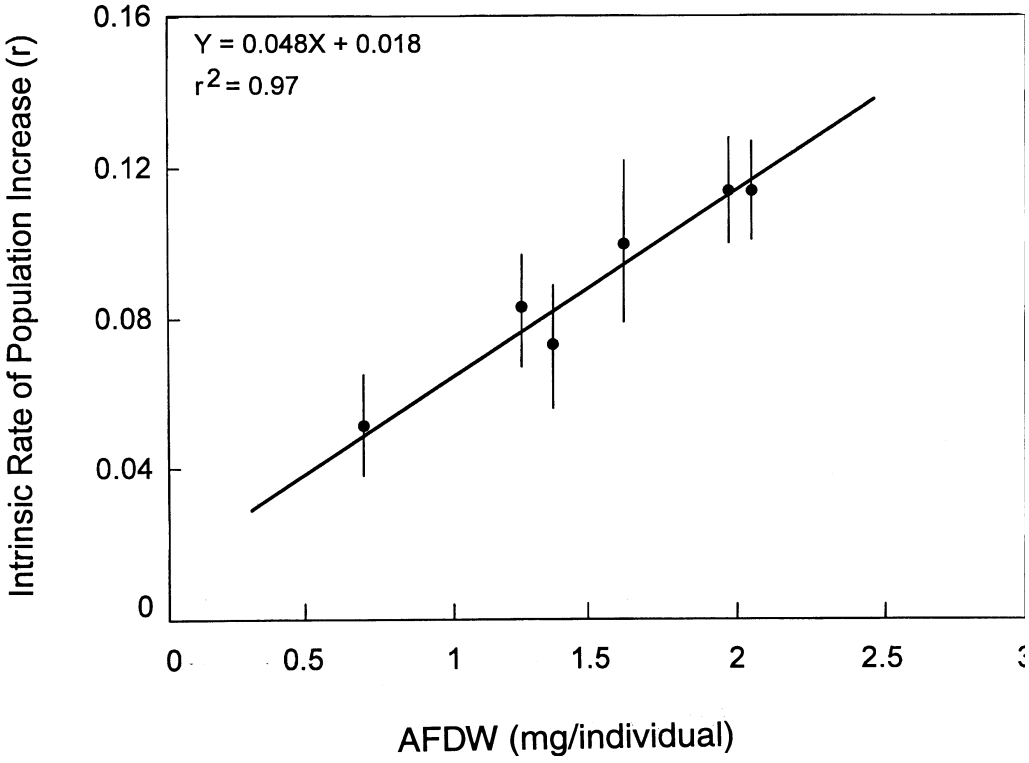


Figure 15.4 Relationship between ash-free dry weight (AFDW) and intrinsic rate of natural increase of *Chironomus tentans.*

15.4.6.4 Relative Endpoint Variability

15.4.6.4.1 Based on coefficient of variation (CV) determined from a control sediment (West Bearskin), the following variability has been documented for the various endpoints in the *C. tentans* life-cycle test (Sibley et al., 1996; Benoit et al., 1997): Survival (<20%), growth as dry weight (<15%), emergence (<30%), reproduction as mean eggs/female (<20%), percent hatch (<10%). Additional information regarding the variation in these endpoints in round-robin testing is presented in Section 17.6.

15.4.6.5 Relative Endpoint Sensitivity

15.4.6.5.1 Measurement of sublethal endpoints (e.g., growth) can often provide unique information in addition to measuring survival. A comparison of lethal and sublethal endpoints relative to toxicity identification is presented in Table 14.4 for *H. azteca*. However, few studies have compared the relative sensitivity of the various endpoints in the *C. tentans* life cycle or in 10-d tests. Sibley et al. (1997a) found that larval *C. tentans* exposed to a gradient of food stress did not experience significant effects on survival, yet did experience a significant reduction in growth and reproduction. Further, the proportion of larvae hatching in this study was high (>80%) and not systematically related to treatment, suggesting that percent hatch may be a relatively insensitive endpoint to sediment-associated contaminants. This is consistent with the findings of another study using zinc-spiked sediments; no effect on embryo viability was observed for those treatments in which egg masses were produced (Sibley et al. 1996). Although the responses observed in the feeding study were not due to a contaminant stressor per se, the sublethal endpoints were clearly better able to discriminate the presence of the stressor than was lethality. Ankley and DeFoe (1998) studied a variety of contaminated sediments and found that the sensitivity of *C. tentans* 10-d tests is greatly increased by measurement of growth in addition to survival. Growth of midge in these 10-d sediment tests was found to be a more sensitive endpoint than survival of *Hyalella azteca*.

15.4.7 Future Research

15.4.7.1 Additional studies using known concentration gradients in sediment, should be conducted to better differentiate the relative sensitivity between lethal and sublethal endpoints and between sublethal endpoints in the long-term *C. tentans* test. Additional studies also are needed to further evaluate the influence of ammonia on long-term exposures with *C. tentans*. Section 1.3.8.5 addresses interpretative guidance for evaluating toxicity associated with ammonia in sediment. Planned water-only toxicity tests with select chemicals (i.e., cadmium, DDD, and fluoranthene) should generate data that can be used to better determine the relative sensitivity of survival, reproduction, and growth endpoints in tests with *C. tentans*. In addition to studies evaluating the relative sensitivity of endpoints, research is also needed to evaluate the ability of these laboratory endpoints to estimate responses of benthic organisms exposed in the field to chemicals in sediments.

Section 16
Data Recording, Data Analysis and Calculations, and Reporting

16.1 Data Recording

16.1.1 Quality assurance project plans with data quality objectives and standard operating procedures should be developed before starting a test. Procedures should be developed by each laboratory to verify and archive data (USEPA, 1994e).

16.1.2 A file should be maintained for each sediment test or group of tests on closely related samples (Section 9). This file should contain a record of the sample chain-of-custody; a copy of the sample log sheet; the original bench sheets for the test organism responses during the sediment test(s); chemical analysis data on the sample(s); control data sheets for reference toxicants; detailed records of the test organisms used in the test(s), such as species, source, age, date of receipt, and other pertinent information relating to their history and health; information on the calibration of equipment and instruments; test conditions used; and results of reference-toxicity tests. Original data sheets should be signed and dated by the laboratory personnel performing the tests. A record of the electronic files of data should also be included in the file.

16.1.3 Example data sheets are included in Appendix D.

16.2 Data Analysis

16.2.1 Statistical methods are used to make inferences about populations, based on samples from those populations. In most sediment toxicity and bioaccumulation tests, test organisms are exposed to chemicals in sediment to estimate the response of the population of laboratory organisms. The organism response to these sediments is usually compared with the response to a control or reference sediment, or in some analyses of bioaccumulation test data, with a fixed standard such as a Food and Drug Administration (FDA) action level. In any toxicity or bioaccumulation test, summary statistics such as means and standard errors for response variables (e.g., survival, chemical concentrations in tissue) should be provided for each treatment (e.g., pore-water concentration, sediment).

16.2.1.1 Types of Data.

16.2.1.1.1 Two types of data can be obtained from sediment toxicity or bioaccumulation tests. The most common endpoint in toxicity testing is mortality, which is a dichotomous or categorical type of data. Other endpoints measured in sublethal evaluations include growth and reproduction (Sections 14 and 15) or tissue concentrations (e.g., in sediment bioaccumulation tests conducted with oligochaetes (Section 13) or with polychaetes and mollusks; USEPA, 1994b). Growth, reproduction, and bioaccumulation endpoints are representative of continuous data.

16.2.1.2 Sediment Testing Scenarios

16.2.1.2.1 Sediment tests are conducted to determine whether contaminants in sediment are harmful to or are bioaccumulated in benthic organisms. Sediment tests are commonly used in studies designed to (1) evaluate dredged material, (2) assess site contamination in the environment (e.g., to rank areas for cleanup), and (3) determine effects of specific contaminants, or combinations of contaminants, through the use of sediment-spiking techniques. Each of these broad study designs has specific statistical design and analytical considerations, which are detailed below.

16.2.1.2.2 Dredged Material Evaluation. In these studies, each site is compared individually with a reference sediment. The statistical procedures appropriate for these studies are generally pairwise comparisons. Additional information on toxicity testing of dredged material and analysis of data from dredged material evaluations is available in USEPA-USACE (1998a).

16.2.1.2.3 Site Assessment of Field Contamination. Surveys of sediment toxicity or bioaccumulation often are included in more comprehensive analyses of biological, chemical, geological, and hydrographic data. Statistical correlation can be improved and costs may be reduced if subsamples are taken simultaneously for sediment toxicity or bioaccumulation tests, chemical analyses, and benthic community structure determinations. There are several statistical approaches to field assessments, each with a specific purpose. If the objective is to compare the response or residue level at all sites individually to a control sediment, then the pairwise comparison approach described below is appropriate. If the objective is to

compare among all sites in the study area, then a multiple comparison procedure that employs an experiment-wise error rate is appropriate. If the objective is to compare among groups of sites, then orthogonal contrasts are a useful data analysis technique.

16.2.1.2.4 Sediment-spiking Experiments. Sediments spiked with known concentrations of chemicals can be used to establish cause-and-effect relationships between chemicals and biological responses. Results of toxicity tests with test materials spiked into sediments at different concentrations may be reported in terms of an LC50, EC50, IC50, NOEC, or LOEC. Results of bioaccumulation tests with either field or spiked samples may be reported in terms of a BSAF (biota sediment accumulation factor; ASTM, 1999c). The statistical approach outlined above for spiked-sediment toxicity tests also applies to the analysis of data from sediment dilution experiments or water-only reference-toxicity tests.

16.2.2 Experimental Design

16.2.2.1 The guidance outlined below on the analysis of sediment toxicity and bioaccumulation test data is adapted from a variety of sources including ASTM (1999c), USEPA (1991a), USEPA (1994a), USEPA (1994b), and USEPA-USACE (1998a). The objectives of a sediment toxicity or bioaccumulation test are to quantify contaminant effects on or accumulation in test organisms exposed to natural or spiked sediments or dredged materials and to determine whether these effects are statistically different from those occurring in a control or reference sediment. Each experiment consists of at least two treatments: the control and one or more test treatment(s). The test treatment(s) consist(s) of the contaminated or potentially contaminated sediment(s). A control sediment is always required to ensure that no contamination is introduced during the experiment setup and that test organisms are healthy. A control sediment is used to judge the acceptability of the test (Tables 11.3, 12.3, 13.4, 14.3, 15.3). Some designs also require a reference sediment that represents an environmental condition or potential treatment effect of interest. Controls are used to evaluate the acceptability of the test and might include a control sediment, a sand substrate (for *C. tentans*; Section 12.2, 15.2), or water-only exposures (for *H. azteca*; Section 14.3.7.8). Testing a reference sediment provides a site-specific basis for evaluating toxicity of the test sediments. Comparisons of test sediments to multiple reference or control sediments representative of the physical characteristics of the test sediment (i.e., grain size, organic carbon) may be useful in these evaluations (Section 2.1.2).

16.2.2.2 Experimental Unit

16.2.2.2.1 During toxicity testing, each test chamber to which a single application of treatment is applied is an experimental unit. During bioaccumulation testing, however, the test organism may be the experimental unit if individual members of the test species are evaluated and they are large enough to provide sufficient biomass for chemical analysis. The important concept is that the treatment (sediment) is applied to each experimental unit as a discrete unit. Experimental units should be independent and should not differ systematically.

16.2.2.3 Replication

16.2.2.3.1 Replication is the assignment of a treatment to more than one experimental unit. The variation among replicates is a measure of the within-treatment variation and provides an estimate of within-treatment error for assessing the significance of observed differences between treatments.

16.2.2.4 Minimum Detectable Difference (MDD)

16.2.2.4.1 As the minimum difference between treatments which the test is required or designed to detect decreases, the number of replicates required to meet a given significance level and power increases. Because no consensus currently exists on what constitutes a biologically acceptable MDD, the appropriate statistical minimum significant difference should be a data quality objective (DQO) established by the individual user (e.g., program considerations) based on their data requirements, the logistics and economics of test design, and the ultimate use of the sediment toxicity or bioaccumulation test results.

16.2.2.5 Minimum Number of Replicates

16.2.2.5.1 Eight replicates are recommended for 10-d freshwater sediment toxicity testing (Section 11 and 12) and five replicates are recommended for 10-d marine testing (USEPA, 1994b). However, four replicates per treatment are the absolute minimum number of replicates for a 10-d sediment toxicity test. A minimum of five replicates per treatment is recommended for bioaccumulation testing (Section 13). It is always prudent to include as many replicates in the test design as are economically and logistically possible. USEPA 10-d sediment toxicity testing methods recommend the use of 10 organisms per replicate for freshwater testing or 20 organisms per replicate for 10-d marine testing. An increase in the number of organisms per replicate in all treatments is allowable only if (1) test performance criteria for the recommended number of replicates are achieved and (2) it can be demonstrated that no change occurs in contaminant availability due to the increased organism loading. See Tables 14.1 and 15.1 for a description of the number of replicates and test organisms/replicate recommended for long-term testing of *Hyalella azteca* or *Chironomus tentans*.

16.2.2.6 Randomization

16.2.2.6.1 Randomization is the unbiased assignment of treatments within a test system and to the exposure chambers ensuring that no treatment is favored and that observations are independent. It is also important to (1) randomly select the organisms (but not the number of organisms) for assignment to the control and test treatments (e.g., a bias in the results may occur if all of the largest animals are placed in the same treatment),

(2) randomize the allocation of sediment (e.g., do not take all the sediment in the top of a jar for the control and the bottom for spiking), and (3) randomize the location of exposure units.

16.2.2.7 Pseudoreplication

16.2.2.7.1 The appropriate assignment of treatments to the replicate exposure chambers is critical to the avoidance of a common error in design and analysis termed "pseudoreplication" (Hurlbert, 1984). Pseudoreplication occurs when inferential statistics are used to test for treatment effects even though the treatments are not replicated or the replicates are not statistically independent (Hurlbert, 1984). The simplest form of pseudoreplication is the treatment of subsamples of the experimental unit as true replicates. For example, two aquaria are prepared, one with control sediment and the other with test sediment, and 10 organisms are placed in each aquarium. Even if each organism is analyzed individually, the 10 organisms only replicate the biological response and do not replicate the treatment (i.e., sediment type). In this case, the experimental unit is the 10 organisms and each organism is a subsample. A less obvious form of pseudoreplication is the potential systematic error due to the physical segregation of exposure chambers by treatment. For example, if all the control exposure chambers are placed in one area of a room and all the test exposure chambers are in another, spatial effects (e.g., different lighting, temperature) could bias the results for one set of treatments. Random physical intermixing of the exposure chambers or randomization of treatment location may be necessary to avoid this type of pseudoreplication. Pseudoreplication can be avoided or reduced by properly identifying the experimental unit, providing replicate experimental units for each treatment, and applying the treatments to each experimental unit in a manner that includes random physical intermixing (interspersion) and independence. However, avoiding pseudoreplication completely may be difficult or impossible given resource constraints.

16.2.2.8 Optimum Design of Experiments

16.2.2.8.1 An optimum design is one which obtains the most precise answer for the least effort. It maximizes or minimizes one of many optimality criteria, which are formal, mathematical expressions of certain properties of the model that are fit to the data. Optimum design of experiments using specific approaches described in Atkinson and Donev (1992) has not been formally applied to sediment testing; however, it might be desirable to use the approaches in experiments. The choice of optimality criterion depends on the objective of the test, and composite criteria can be used when a test has more than one goal. A design is optimum only for a specific model, so it is necessary to know beforehand which models might be used (Atkinson and Donev, 1992).

16.2.2.9 Compositing Samples

16.2.2.9.1 Decisions regarding compositing of samples depend on the objective of the test. Compositing is used primarily in bioaccumulation experiments when the biomass of an individual organism is insufficient for chemical analysis. Compositing consists of combining samples (e.g., organisms, sediment) and chemically analyzing the mixture rather than the individual samples. The chemical analysis of the mixture provides an estimate of the average concentration of the individual samples making up the composite. Compositing also may be used when the cost of analysis is high. Each organism or sediment sample added to the composite should be of equal size (i.e., wet weight) and the composite should be completely homogenized before taking a sample for chemical analysis. If compositing is performed in this manner, the value obtained from the analysis of the composite is the same as the average obtained from analyzing each individual sample (within any sampling and analytical errors). If true replicate composites (not subsample composites) are made, the variance of the replicates will be less than the variance of the individual samples, providing a more precise estimate of the mean value. This increases the power of a test between means of composites over a test between means of individuals or samples for a given number of samples analyzed. If compositing reduces the actual number of replicates, however, the power of the test will also be reduced. If composites are made of individuals or samples varying in size, the value of the composite and the mean of the individual organisms or sediment samples are no longer equivalent. The variance of the replicate composites will increase, decreasing the power of any test between means. In extreme cases, the variance of the composites can exceed the population variance (Tetra Tech, 1986). Therefore, it is important to keep the individuals or sediment samples comprising the composite equivalent in size. If sample sizes vary, consult the tables in Schaeffer and Janardan (1978) to determine if replicate composite variances will be higher than individual sample variances, which would make compositing inappropriate.

16.2.3 Hypothesis Testing and Power

16.2.3.1 The purpose of a toxicity or bioaccumulation test is to determine if the biological response to a treatment sample differs from the response to a control sample. Figure 16.1 presents the possible outcomes and decisions that can be reached in a statistical test of such a hypothesis. The null hypothesis is that no difference exists among the mean control and treatment responses. The alternative hypothesis of greatest interest in sediment tests is that the treatments are toxic, or contain concentrations of bioaccumulatable compounds, relative to the control or reference sediment.

16.2.3.2 Statistical tests of hypotheses can be designed to control for the chances of making incorrect decisions. In Figure 16.1, alpha (α) represents the probability of making a Type I statistical error. A Type I statistical error in this testing situation results from the false conclusion

Decision	TR =Control	TR > Control
TR =Control	Correct 1 - α	Type II Error β
TR > Control	Type I Error α	Correct 1 - β (Power)

Treatment response (TR), Alpha (α) represents the probability of making a Type I statistical error (false positive); beta (β) represents the probability of making a Type II statistical error (false negative).

Figure 16.1 Treatment response for a Type I and Type II error.

that the treated sample is toxic or contains chemical residues not found in the control or reference sample. Beta (β) represents the probability of making a Type II statistical error, or the likelihood that one erroneously concludes there are no differences among the mean responses in the treatment, control or reference samples. Traditionally, acceptable values for α have ranged from 0.1 to 0.01 with 0.05 or 5% used most commonly. This choice should depend upon the consequences of making a Type I error. Historically, having chosen α, environmental researchers have ignored β and the associated power of the test (1-β).

16.2.3.3 Fairweather (1991) presents a review of the need for, and the practical implications of, conducting power analyses in environmental monitoring studies. This review also includes a comprehensive bibliography of recent publications on the need for, and use of, power analyses in environmental study design and data analysis. The consequences of a Type II statistical error in environmental studies should never be ignored and may, in fact, be one of the most important criteria to consider in experimental designs and data analyses that include statistical hypothesis testing. To paraphrase Fairweather (1991), "The commitment of time, energy and people to a false positive (a Type I error) will only continue until the mistake is discovered. In contrast, the cost of a false negative (a Type II error) will have both short- and long-term costs (e.g., ensuing environmental degradation and the eventual cost of its rectification)."

16.2.3.4 The critical components of the experimental design associated with the testing of hypotheses outlined above are (1) the required MDD between the treatment and control or reference responses, (2) the variance among treatment and control replicate experimental units, (3) the number of replicate units for the treatment and control samples, (4) the number of animals exposed within a replicate exposure chamber, and (5) the selected probabilities of Type I (α) and Type II (β) errors.

16.2.3.5 Sample size or number of replicates may be fixed due to cost or space considerations or may be varied to achieve *a priori* probabilities of α and β. The MDD should be established ahead of time based upon biological and program considerations. The investigator has little control of the variance among replicate exposure chambers. However, this variance component can be minimized by selecting test organisms that are as biologically similar as possible and maintaining test conditions within prescribed quality control (QC) limits.

16.2.3.6 The MDD is expressed as a percentage change from the mean control response. To test the equality of the control and treatment responses, a two-sample *t* test with its associated assumptions is the appropriate parametric analysis. If the desired MDD, the number of replicates per treatment, the number of organisms per replicate and an estimate of typical among replicate variability, such as the coefficient of variation (CV) from a control sample, are available, it is possible to use a graphical approach as in Figure 16.2 to determine how likely it is that a 20% reduction will be detected in the treatment response relative to the control response. The CV is defined as 100% x (standard deviation divided by the mean). In a test design with 8 replicates per treatment and with an α level of 0.05, high power (i.e., >0.8) to detect a 20% reduction from the control mean occurs only if the CV is 15% or less (Figure 16.2). The choice of these variables also affects the power of the test. If 5 replicates are used per treatment (Figure 16.3), the CV needs to be 10% or lower to detect a 20% reduction in response relative to the control mean with a power of 90%.

16.2.3.7 Relaxing the α level of a statistical test increases the power of the test. Figure 16.4 duplicates Figure 16.2 except that α is 0.10 instead of 0.05. Selection of the appropriate α level of a test is a function of the costs associated with making Type I and II statistical errors. Evaluation of Figure 16.2 illustrates that with a CV of 15% and an α level of 0.05, there is an 80% probability (power) of detecting a 20% reduction in the mean treatment response relative to the control mean. However, if α is set at 0.10 (Figure 16.4) and the CV remains at 15%, then there is a 90% probability (power) of detecting a 20% reduction relative to the control mean. The latter example would be preferable if an environmentally conservative analysis and interpretation of the data is desirable.

16.2.3.8 Increasing the number of replicates per treatment will increase the power to detect a 20% reduction in treatment response relative to the control mean (Figure 16.5). Note, however, that for less than 8 replicates per treatment it is difficult to have high power (i.e., >0.80) unless the CV is less than 15%. If space or cost limit the number of replicates to fewer than 8 per treatment, then it may be necessary to find ways to reduce the among replicate variability and consequently the CV. Options that are available to increase the power of the test include selecting more uniform organisms to reduce biological variability or increasing the α level of the test. For CVs in the range of 30% to 40%, even 8 replicates per treatment is inadequate to detect small reductions (<20%) in response relative to the control mean.

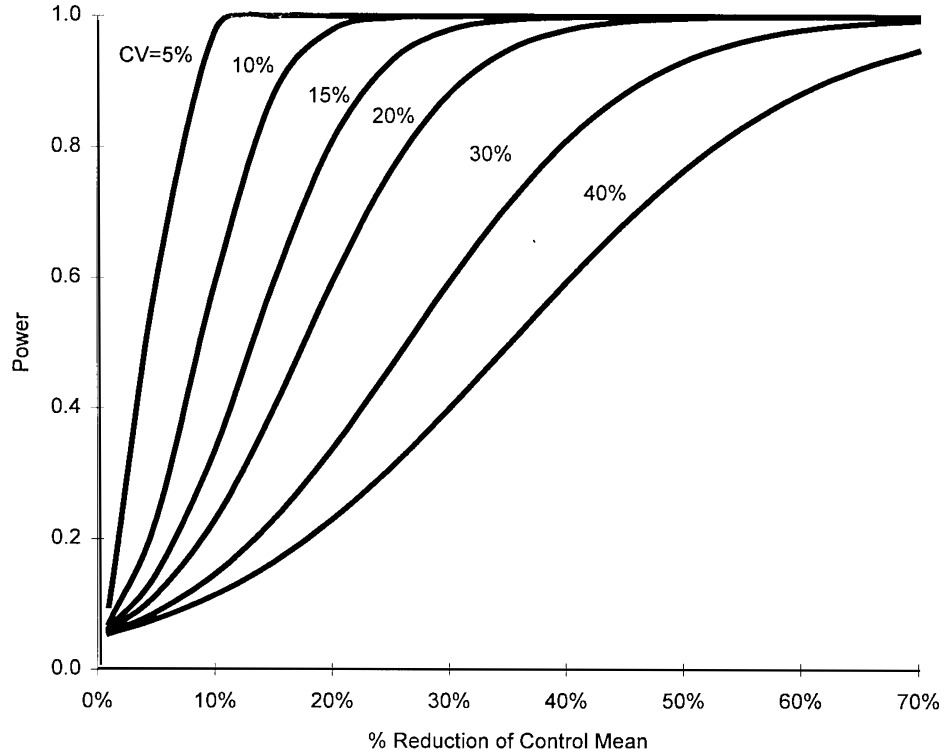


Figure 16.2 **Power of the test vs. percent reduction in treatment response relative to the control mean at various CVs (8 replicates, alpha = 0.05 [one-tailed]).**

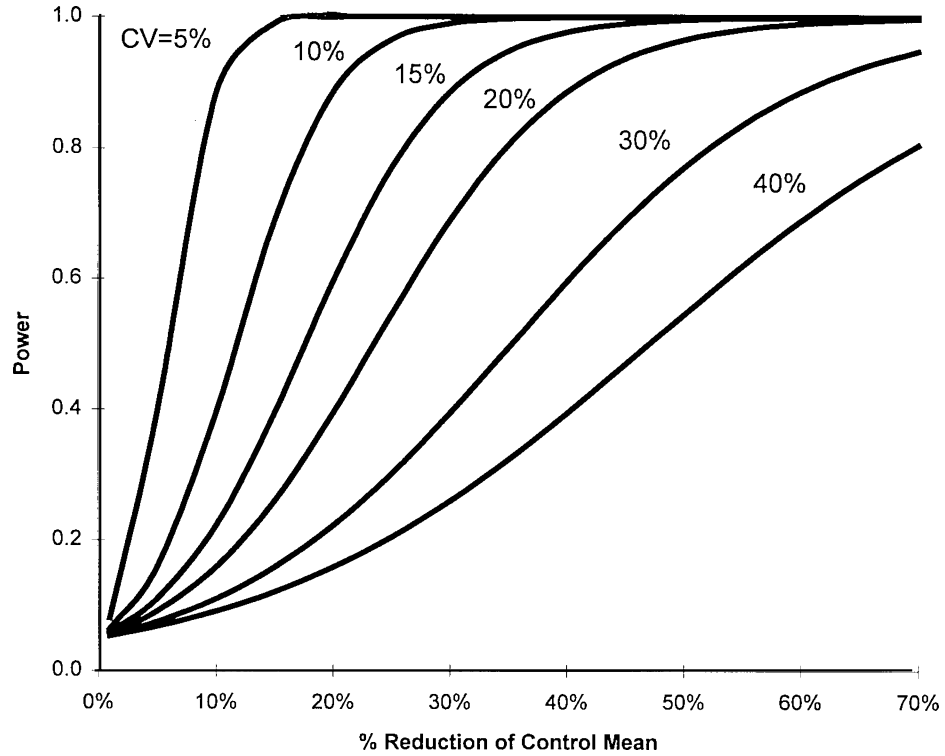


Figure 16.3 **Power of the test vs. percent reduction in treatment response relative to the control mean at various CVs (5 replicates, alpha = 0.05 [one-tailed]).**

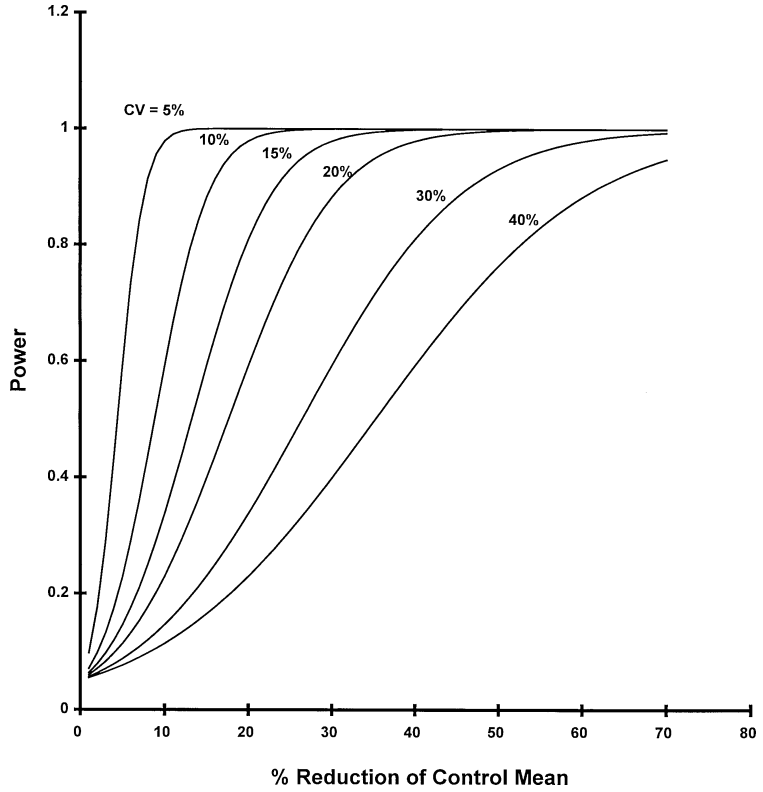


Figure 16.4 Power of the test vs. percent reduction in treatment response relative to the control mean at various CVs (8 replicates, alpha = 0.10 [one-tailed]).

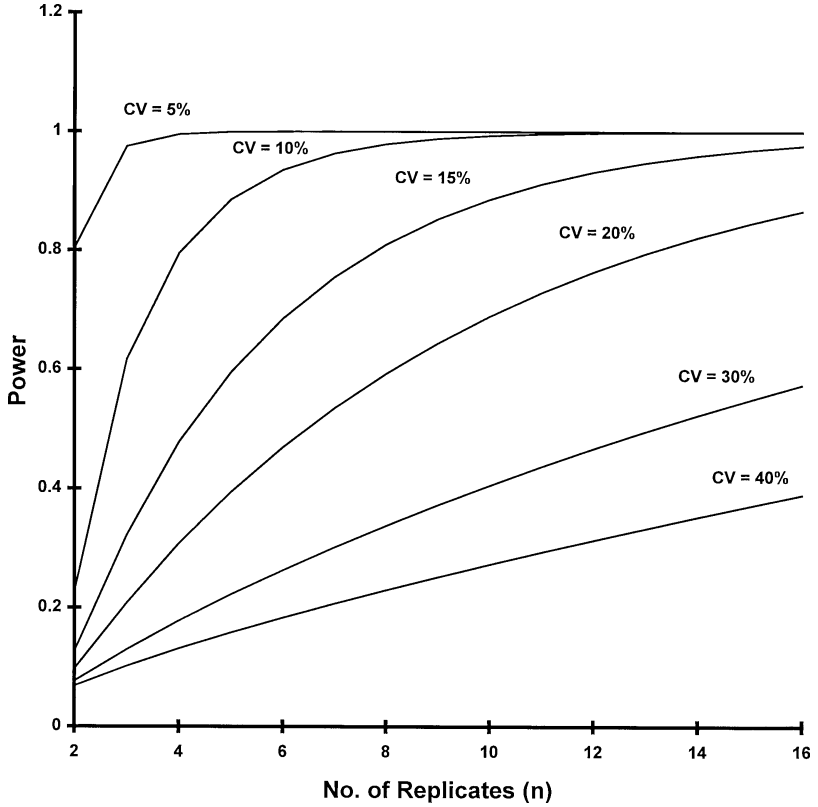


Figure 16.5 Effect of CV and number of replicates on the power to detect a 20% decrease in treatment response relative to the control mean (alpha = 0.05 [one-tailed]).

16.2.3.9 The effect of the choice of α and β on number of replicates for various CVs, assuming the combined total probability of Type I and Type II statistical errors is fixed at 0.25, is illustrated in Figure 16.6. An α of 0.10 therefore establishes a β of 0.15. In Figure 16.6, if $\alpha = \beta = 0.125$, the number of replicates required to detect a difference of 20% relative to the control is at a minimum. As α or β decrease, the number of replicates required to detect the same 20% difference relative to the control increases. However, the curves are relatively flat over the range of 0.05 to 0.20, and their shape will change dramatically if the combined total $\alpha + \beta$ is changed. Limiting the total of $\alpha + \beta$ to 0.10 greatly increases the number of replicates necessary to detect a preselected percentage reduction in mean treatment response relative to the control mean.

16.2.4 Comparing Means

16.2.4.1 Figure 16.7 outlines a decision tree for analysis of survival, growth, or reproduction data subjected to hypothesis testing. In the tests described herein, samples or observations refer to replicates of treatments. Sample size n is the number of replicates (i.e., exposure chambers) in an individual treatment, not the number of organisms in an exposure chamber. Overall sample size N is the combined total number of replicates in all treatments. The statistical methods discussed in this section are described in general statistics texts such as Steel and Torrie (1980), Sokal and Rohlf (1981), Dixon and Massey (1983), Zar (1984), and Snedecor and Cochran (1989). It is recommended that users of this manual have at least one of these texts and associated statistical tables on hand. A nonparametric statistics text such as Conover (1980) might also be helpful.

16.2.4.2 Mean

16.2.4.2.1 The sample mean ($\bar{x}$) is the average value, or $\Sigma x_i/n$ where

n = number of observations (replicates)

x_i = ith observation

Σx_i = every x summed = $x_1 + x_2 + x_3 + \ldots + x_n$

16.2.4.3 Standard Deviation

16.2.4.3.1 The sample standard deviation (s) is a measure of the variation of the data around the mean and is equivalent to $\sqrt{s^2}$. The sample variance, s^2, is given by the following "machine" or "calculation" formula:

$$s^2 = \frac{\sum_{i=1}^{n} x_i^2 - (\bar{x})^2}{n-1}$$

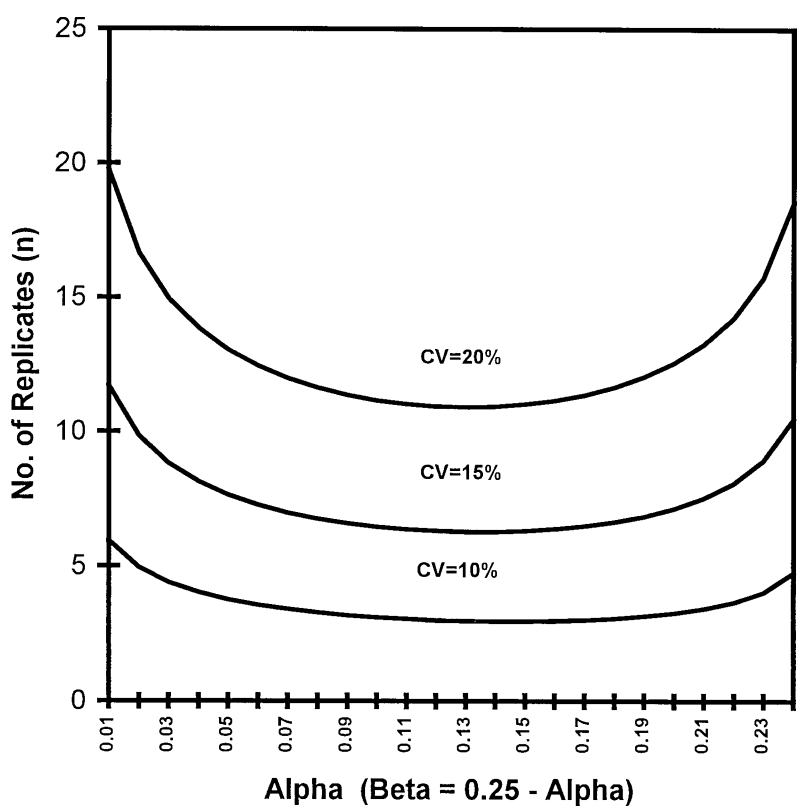


Figure 16.6 Effect of alpha and beta on the number of replicates at various CVs (assuming combined alpha + beta = 0.25).

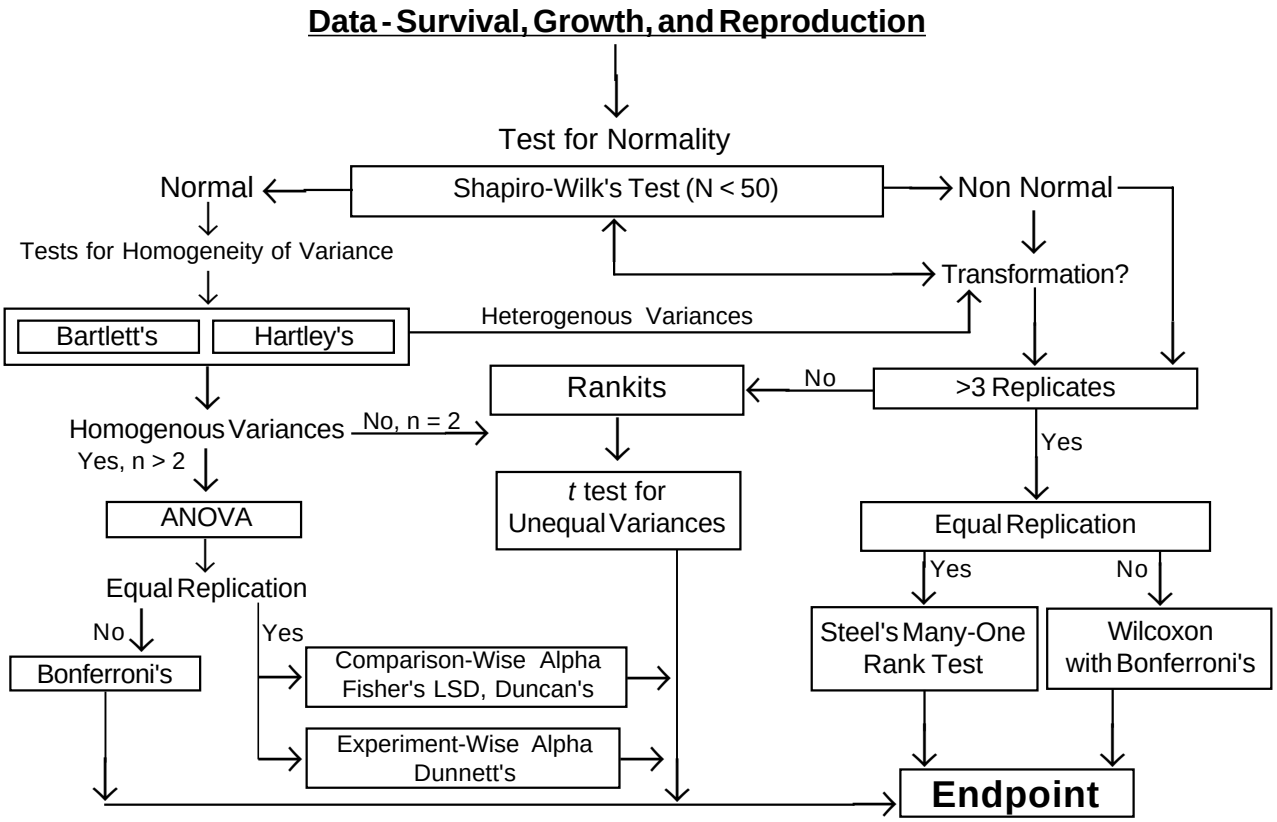


Figure 16.7 Decision tree for analysis of survival, growth, and reproduction data subjected to hypothesis testing.

16.2.4.4 Standard Error of the Mean

16.2.4.4.1 The standard error of the mean (SE, or $s/\sqrt{n}$) estimates variation among sample means rather than among individual values. The SE is an estimate of the standard deviation among means that would be obtained from several samples of n observations each. Most of the statistical tests in this manual compare means with other means (e.g., dredged sediment mean with reference mean) or with a fixed standard (e.g., FDA action level; ASTM, 1999c). Therefore, the "natural" or "random" variation of sample means (estimated by SE), rather than the variation among individual observations (estimated by s), is required for the tests.

16.2.4.5 Tests of Assumptions

16.2.4.5.1 In general, parametric statistical analyses such as *t* tests and analysis of variance are appropriate only if (1) there are independent, replicate experimental units for each treatment, (2) the observations within each treatment follow a normal distribution, and (3) variances for both treatments are equal or similar. The first assumption is an essential component of experimental design. The second and third assumptions can be tested using the data obtained from the experiment. Therefore, before conducting statistical analyses, tests for normality and equality of variances should be performed.

16.2.4.5.2 Outliers. Extreme values and systematic departures from a normal distribution (e.g., a log-normal distribution) are the most common causes of departures from normality or equality of variances. An outlier is an inconsistent or questionable data point that appears unrepresentative of the general trend exhibited by the majority of the data. Outliers may be detected by tabulation of the data, by plotting, or by analysis of residuals. An explanation should be sought for any questionable data points. Without an explanation, data points should only be discarded with extreme caution. If there is no explanation, the analysis should be performed both with and without the outlier, and the results of both analyses should be reported. An appropriate transformation, such as the arc sine-square root transformation, will normalize many distributions (USEPA, 1985). Problems with outliers can usually be solved only by using nonparametric tests, but careful laboratory practices can reduce the frequency of outliers.

16.2.4.5.3 Tests for Normality. The most commonly used test for normality for small sample sizes (N<50) is the Shapiro-Wilk's test. This test determines if residuals are normally distributed. Residuals are the differences between individual observations and the treatment mean. Residuals, rather than raw observations, are tested because subtracting the treatment mean removes any differences among treatments. This scales the observations so that the mean of residuals for each treatment and over

all treatments is zero. The Shapiro-Wilk's test provides a test statistic W, which is compared to values of W expected from a normal distribution. W will generally vary between 0.3 and 1.0, with lower values indicating greater departure from normality. Because normality is desired, one looks for a high value of W with an associated probability greater than the pre-specified α level.

16.2.4.5.3.1 Table 16.1 provides α levels to determine whether departures from normality are significant. Normality should be rejected when the probability associated with W (or other normality test statistic) is less than α for the appropriate total number of replicates (N) and design. A balanced design means that all treatments have an equal number (n) of replicate exposure chambers. A design is considered unbalanced when the treatment with the largest number of replicates (n_{max}) has at least twice as many replicates as the treatment with the fewest replicates (n_{min}). Note that higher α levels are used when the number of replicates is small, or when the design is unbalanced, because these are the cases in which departures from normality have the greatest effects on *t* tests and other parametric comparisons. If data fail the test for normality, even after transformation, nonparametric tests should be used for additional analyses (See Section 16.2.4.8 and Figure 16.7).

16.2.4.5.3.2 Tables of quantiles of W can be found in Shapiro and Wilk (1965), Gill (1978), Conover (1980), USEPA (1989c) and other statistical texts. These references also provide methods of calculating W, although the calculations can be tedious. For that reason, commonly available computer programs or statistical packages are preferred for the calculation of W.

16.2.4.5.4 Tests for Homogeneity of Variances. There are a number of tests for equality of variances. Some of these tests are sensitive to departures from normality, which is why a test for normality should be performed first. Bartlett's test or other tests such as Levene's test or Cochran's test (Winer, 1971; Snedecor and Cochran, 1989) all have similar power for small, equal sample sizes (n=5) (Conover et al., 1981). The data must be normally distributed for Bartlett's test. Many software packages for *t* tests and analysis of variance (ANOVA) provide at least one of the tests.

Table 16.1 Suggested α Levels to Use for Tests of Assumptions

Test	Number of Observations[1]	α When Design Is Balanced	Unbalanced[2]
Normality	N = 2 to 9	0.10	0.25
	N = 10 to 19	0.05	0.10
	N = 20 or more	0.01	0.05
Equality of variances	n = 2 to 9	0.10	0.25
	n = 10 or more	0.05	0.10

[1] N = total number of observations (replicates) in all treatments combined; n = number of observations (replicates) in an individual treatment

[2] $n_{max} \bullet 2\ n_{min}$

16.2.4.5.4.1 If no tests for equality of variances are included in the available statistical software, Hartley's F_{max} can easily be calculated:

$$F_{max} = (\text{ larger of } s_1^2\ ,\ s_2^2\ ) / (\text{ smaller of } s_1^2\ ,\ s_2^2\ )$$

When F_{max} is large, the hypothesis of equal variances is more likely to be rejected. F_{max} is a two-tailed test because it does not matter which variance is expected to be larger. Some statistical texts provide critical values of F_{max} (Winer, 1971; Gill, 1978; Rohlf and Sokal, 1981).

16.2.4.5.4.2 Levels of α for tests of equality of variances are provided in Table 16.1. These levels depend upon number of replicates in a treatment (n) and allotment of replicates among treatments. Relatively high α's (i.e., $\geq$0.10) are recommended because the power of the above tests for equality of variances is rather low (about 0.3) when n is small. Equality of variances is rejected if the probability associated with the test statistic is less than the appropriate α.

16.2.4.6 Transformations of the Data

16.2.4.6.1 When the assumptions of normality or homogeneity of variance are not met, transformations of the data may remedy the problem, so that the data can be analyzed by parametric procedures, rather than by a nonparametric technique. The first step in these analyses is to transform the responses, expressed as the proportion surviving, by the arc sine-square root transformation. The arc sine-square root transformation is commonly used on proportionality data to stabilize the variance and satisfy the normality requirement. If the data do not meet the assumption of normality and there are four or more replicates per group, then the nonparametric test, Wilcoxon Rank Sum test, can be used to analyze the data. If the data meet the assumption of normality, Bartlett's test or Hartley's *F* test for equality of variances is used to test the homogeneity of variance assumption. Failure of the homogeneity of variance assumption leads to the use of a modified *t* test, and the degrees of freedom for the test are adjusted.

16.2.4.6.2 The arc sine-square root transformation consists of determining the angle (in radians) represented by a sine value. In this transformation, the proportion surviving is taken as the sine value, the square root of the sine value is calculated, and the angle (in radians) for the square root of the sine value is determined. When the proportion surviving is 0 or 1, a special modification of the transformation should be used (Bartlett, 1937). An example of the arc sine-square root transformation and modification are provided below.

1. Calculate the response proportion (RP) for each replicate within a group, where

   RP = (number of surviving organisms)/(number exposed)

2. Transform each RP to arc sine, as follows:

   a. For RPs greater than zero or less than one:

   Angle (in radians) = *arc sine* $\sqrt{(RP)}$

   b. Modification of the arc sine when RP = 0.

   Angle (in radians) = *arc sine* $\sqrt{\frac{1}{4n}}$

   where n = number of animals/treatment rep.

   c. Modification of the arc sine when RP = 1.0.

   Angle = 1.5708 radians-(radians for RP = 0)

16.2.4.7 Two Sample Comparisons (N=2)

16.2.4.7.1 The true population mean (μ) and standard deviation (σ) are known only after sampling the entire population. In most cases, samples are taken randomly from the population, and the s calculated from those samples is only an estimate of σ. Student's *t*-values account for this uncertainty. The degrees of freedom for the test, which are defined as the sample size minus one (n-1), should be used to obtain the correct *t*-value. Student's *t*-values decrease with increasing sample size because larger samples provide a more precise estimate of μ and σ.

16.2.4.7.2 When using a *t* table, it is crucial to determine whether the table is based on one-tailed probabilities or two-tailed probabilities. In formulating a statistical hypothesis, the alternative hypothesis can be one-sided (one-tailed test) or two-sided (two-tailed test). The null hypothesis (H_0) is always that the two values being analyzed are equal. A one-sided alternative hypothesis (H_a) is that there is a specified relationship between the two values (e.g., one value is greater than the other) versus a two-sided alternative hypothesis (H_a) which is that the two values are simply different (i.e., either larger or smaller). A one-tailed test is used when there is an *a priori* reason to test for a specific relationship between two means, such as the alternative hypothesis that the treatment mortality or tissue residue is greater than the control mortality or tissue residue. In contrast, the two-tailed test is used when the direction of the difference is not important or cannot be assumed before testing.

16.2.4.7.3 Since control organism mortality or tissue residues and sediment chemical concentrations are presumed lower than reference or treatment sediment values, conducting one-tailed tests is recommended in most cases. For the same number of replicates, one-tailed tests are more likely to detect statistically significant differences between treatments (e.g., have a greater power) than are two-tailed tests. This is a critical consideration when dealing with a small number of replicates (such as 8/treatment). The other alternative for increasing statistical power is to increase the number of replicates, which increases the cost of the test.

16.2.4.7.4 There are cases when a one-tailed test is inappropriate. When no *a priori* assumption can be made as to how the values vary in relationship to one another, a two-tailed test should be used. An example of an alternative two-sided hypothesis is that the reference sediment total organic carbon (TOC) content is different (greater or lesser) from the control sediment TOC. A two-tailed test should also be used when comparing tissue residues among different species exposed to the same sediment and when comparing bioaccumulation factors (BAFs) or biota-sediment accumulation factors (BSAFs).

16.2.4.7.5 The *t*-value for a one-tailed probability can be found in a two-tailed table by looking up *t* under the column for twice the desired one-tailed probability. For example, the one-tailed *t*-value for $\alpha = 0.05$ and df = 20 is 1.725, and is found in a two-tailed table using the column for $\alpha = 0.10$.

16.2.4.7.6 The usual statistical test for comparing two independent samples is the two-sample *t* test (Snedecor and Cochran, 1989). The *t*-statistic for testing the equality of means $\overline{x}_1$ and $\overline{x}_2$ from two independent samples with n_1 and n_2 replicates and unequal variances is

$$t = (\overline{x}_1 - \overline{x}_2) / \sqrt{s_1^2 / n_1 + s_2^2 / n_2}$$

where s_1^2 and s_2^2 are the sample variances of the two groups. Although the equation assumes that the variances of the two groups are unequal, it is equally useful for situations in which the variances of the two groups are equal. This statistic is compared with the Student's *t* distribution with degrees of freedom (df) given by Satterthwaite's (1946) approximation:

$$df = \frac{(s_1^2 / n_1 + s_2^2 / n_2)^2}{(s_1^2 / n_1)^2 / (n_1 - 1) + (s_2^2 / n_2)^2 / (n_2 - 1)}$$

This formula can result in fractional degrees of freedom, in which case one should round the degree of freedom down to the nearest integer in order to use a *t* table. Using this approach, the degrees of freedom for this test will be less than the degrees of freedom for a *t* test assuming equal variances. If there are unequal numbers of replicates in the treatments, the *t* test with Bonferroni's adjustment can be used for data analysis (USEPA, 1994c; USEPA, 1994d). When variances are equal, an *F* test for equality is unnecessary.

16.2.4.8 Nonparametric Tests

16.2.4.8.1 Tests such as the *t* test, which analyze the original or transformed data and which rely on the properties of the normal distribution, are referred to as parametric tests. Nonparametric tests, which do not require normally distributed data, analyze the ranks of data and generally compare medians rather than means. The me-

dian of a sample is the middle or 50th percentile observation when the data are ranked from smallest to largest. In many cases, nonparametric tests can be performed simply by converting the data to ranks or normalized ranks (rankits) and conducting the usual parametric test procedures on the ranks or rankits.

16.2.4.8.2 Nonparametric tests are useful because of their generality, but have less statistical power than corresponding parametric tests when the parametric test assumptions are met. If parametric tests are not appropriate for comparisons because the normality assumption is not met, data should be converted to normalized ranks (rankits). Rankits are simply the z-scores expected for the rank in a normal distribution. Thus, using rankits imposes a normal distribution over all the data, although not necessarily within each treatment. Rankits can be obtained by ranking the data, then converting the ranks to rankits using the following formula:

$$rankit = z_{[(rank - 0.375) / (N + 0.25)]}$$

where z is the normal deviate and N is the total number of observations. Alternatively, rankits may be obtained from standard statistical tables such as Rohlf and Sokal (1981).

16.2.4.8.3 If normalized ranks are calculated, the ranks should be converted to rankits using the formula above. In comparisons involving only two treatments (N=2), there is no need to test assumptions on the rankits or ranks; simply proceed with a one-tailed *t* test for unequal variances using the rankits or ranks.

16.2.4.9 Analysis of Variance (N>2)

16.2.4.9.1 Some experiments are set up to compare more than one treatment with a control, whereas others may also be interested in comparing the treatments with one another. The basic design of these experiments is the same as for experiments evaluating pairwise comparisons. After the applicable comparisons are determined, the data must be tested for normality to determine whether parametric statistics are appropriate and whether the variances of the treatments are equal. If normality of the data and equal variances are established, then an analysis of variance (ANOVA) may be performed to address the hypothesis that all the treatments, including the control, are equal. If normality or equality of variance are not established, then transformations of the data might be appropriate, or nonparametric statistics can be used to test for equal means. Tests for normality of the data should be performed on the treatment residuals. A residual is defined as the observed value minus the treatment mean, that is, $r_{ik} = o_{ik}$ - (k^{th} treatment mean). Pooling residuals provides an adequate sample size to test the data for normality.

16.2.4.9.2 The variances of the treatments should also be tested for equality. Currently there is no easy way to test for equality of the treatment means using analysis of variance if the variances are not equal. In a toxicity test with several treatments, one treatment may have 100% mortality in all of its replicates, or the control treatment may have 100% survival in all of its replicates. These responses result in 0 variance for a treatment that results in a rejection of equality of variance in these cases. No transformation will change this outcome. In this case, the replicate responses for the treatment with 0 variance should be removed before testing for equality of variances. Only those treatments that do not have 0 replicate variance should be used in the ANOVA to get an estimate of the within treatment variance. After a variance estimate is obtained, the means of the treatments with 0 variance can be tested against the other treatment means using the appropriate mean comparison. Equality of variances among the treatments can be evaluated with the Hartley F_{max} test or Bartlett's test. The option of using nonparametric statistics on the entire set of data is also an alternative.

16.2.4.9.3 If the data are not normally distributed or the variances among treatments are not homogeneous, even after data transformation, nonparametric analyses are appropriate. If there are four or more replicates per treatment and the number of replicates per treatment is equal, the data can be analyzed with Steel's Many-One Rank test. Unequal replication among treatments requires data analysis with the Wilcoxon Rank Sum test with Bonferroni's adjustment. Steel's Many-One Rank test is a nonparametric test for comparing treatments with a control. This test is an alternative to the Dunnett's test, and may be applied to data when the normality assumption has not been met. Steel's test requires equal variances across treatments and the control, but is thought to be fairly insensitive to deviations from this condition (USEPA, 1991a). Wilcoxon's Rank Sum test is a nonparametric test to be used as an alternative to the Steel's test when the number of replicates are not the same within each treatment. A Bonferroni's adjustment of the pairwise error rate for comparison of each treatment versus the control is used to set an upper bound of alpha on the overall error rate. This is in contrast to the Steel's test with a fixed overall error rate for alpha. Thus, Steel's test is a more powerful test (USEPA, 1991a).

16.2.4.9.4 Different mean comparison tests are used depending on whether an α percent comparison-wise error rate or an α percent experiment-wise error rate is desired. The choice of a comparison-wise or experiment-wise error rate depends on whether a decision is based on a pairwise comparison (comparison-wise) or from a set of comparisons (experiment-wise). For example, a comparison-wise error rate would be used for deciding which stations along a gradient were acceptable or not acceptable relative to a control or reference sediment. Each individual comparison is performed independently at a smaller α (than that used in an experiment-wise comparison), such that the probability of making a Type I error in the entire series of comparisons is not greater than the chosen experiment-wise α level of the test. This results in a more conservative test when comparing any particular sample to the control or reference. However, if several samples were taken from the same area and the decision to accept or reject the area was based upon all comparisons

with a reference, then an experiment-wise error rate should be used. When an experiment-wise error rate is used, the power to detect real differences between any two means decreases as a function of the number of treatment means being compared to the control treatment.

16.2.4.9.5 The recommended procedure for pairwise comparisons that have a comparison-wise α error rate and equal replication is to do an ANOVA followed by a one-sided Fisher's Least Significant Difference (LSD) test (Steel and Torrie, 1980). A Duncan's mean comparison test should give results similar to the LSD. If the treatments do not contain equal numbers of replicates, the appropriate analysis is the *t* test with Bonferroni's adjustment. For comparisons that maintain an experiment-wise α error rate, Dunnett's test is recommended for comparisons with the control.

16.2.4.9.6 Dunnett's test has an overall error rate of α, which accounts for the multiple comparisons with the control. Dunnett's procedure uses a pooled estimate of the variance, which is equal to the error value calculated in an ANOVA.

16.2.4.9.7 To perform the individual comparisons, calculate the *t* statistic for each treatment and control combination, as follows:

$$t_i = \frac{(\overline{Y_1} - \overline{Y_i})}{S_w \sqrt{(1/n_1) + (1/n_i)}}$$

where $\overline{Y}_i$ = mean for each treatment

$\overline{Y}_1$ = mean for the control

S_w = square root of the within mean square

n_1 = number of replicates in the control

n_i = number of replicates for treatment "i"

To quantify the sensitivity of the Dunnett's test, the minimum significant difference (MSD=MDD) may be calculated with the following formula:

$$MSD = d\ S_w \sqrt{(1/n_1) + (1/n)}$$

where d = Critical value for the Dunnett's Procedure

S_w = The square root of the within mean square

n = The number of replicates per treatment, assuming an equal number of replicates at all treatment concentrations

n_1 = Number of replicates in the control

16.2.5 Methods for Calculating LC50s, EC50s, and ICps

16.2.5.1 Figure 16.8 outlines a decision tree for analysis of point estimate data. USEPA manuals (USEPA, 1991a; USEPA, 1994c; USEPA, 1994d) discuss in detail the mechanics of calculating LC50 (or EC50) or ICp values using the most current methods. The most commonly used methods are the Graphical, Probit, trimmed Spearman-Karber and the Linear Interpolation Methods. Methods for evaluating point estimate data using logistic regression are outlined in Snedecor and Cochran (1989). In general, results from these methods should yield similar estimates. Each method is outlined below, and recommendations are presented for the use of each method.

16.2.5.2 Data for at least five test concentrations and the control should be available to calculate an LC50, although each method can be used with fewer concentrations. Survival in the lowest concentration must be at least 50%, and an LC50 should not be calculated unless at least 50% of the organisms die in at least one of the serial dilutions. When less than 50% mortality occurs in the highest test concentration, the LC50 is expressed as greater than the highest test concentration.

16.2.5.3 Due to the intensive nature of the calculations for the estimated LC50 and associated 95% confidence interval using most of the following methods, it is recommended that the data be analyzed with the aid of computer software. Computer programs to estimate the LC50 or ICp values and associated 95% confidence intervals using the methods discussed below (except for the Graphical Method) were developed by USEPA and can be obtained by sending a diskette with a written request to USEPA, National Exposure Research Laboratory, 26 W.

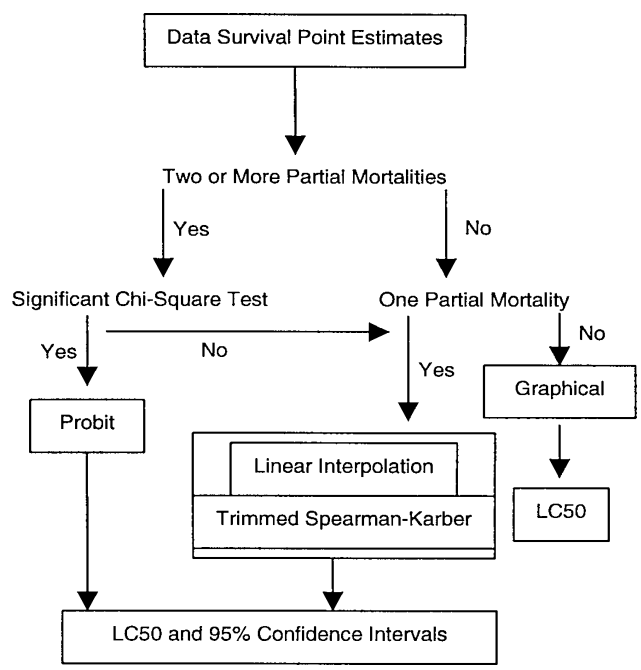


Figure 16.8 Decision tree for analysis of point estimate data.

Martin Luther King Drive, Cincinnati, OH 45268 or call 513/569-7076.

16.2.5.4 Graphical Method

16.2.5.4.1 This procedure estimates an LC50 (or EC50) by linearly interpolating between points of a plot of observed percentage mortality versus the base 10 logarithm ($\log_{10}$) of treatment concentration. The only requirement for its use is that treatment mortalities bracket 50%.

16.2.5.4.2 For an analysis using the Graphical Method, the data should first be smoothed and adjusted for mortality in the control replicates. The procedure for smoothing and adjusting the data is detailed in the following steps: Let p_0, p_1, ..., p_k denote the observed proportion mortalities for the control and the k treatments. The first step is to smooth the p_i if they do not satisfy p_0 - p_i - ... - p_k. The smoothing process replaces any adjacent p_i's that do not conform to p_0 - p_1- ... - p_k with their average. For example, if p_i is less than p_{i-1}, then

$$p_{i-1}^s = p_i^s = (p_i + p_{i-1})/2$$

where p_i^s = the smoothed observed proportion mortality for concentration *i.*

Adjust the smoothed observed proportion mortality in each treatment for mortality in the control group using Abbott's formula (Finney, 1971). The adjustment takes the form:

$$p_i^a = (p_i^s - p_0^s)/(1 - p_0^s)$$

where p_0^s = the smoothed observed proportion mortality for the control

p_i^s = the smoothed observed proportion mortality for concentration *i.*

16.2.5.5 The Probit Method

16.2.5.5.1 This method is a parametric statistical procedure for estimating the LC50 (or EC50) and the associated 95% confidence interval (Finney, 1971). The analysis consists of transforming the observed proportion mortalities with a Probit transformation, and transforming the treatment concentrations to $\log_{10}$. Given the assumption of normality for the $\log_{10}$ of the tolerances, the relationship between the transformed variables mentioned above is about linear. This relationship allows estimation of linear regression parameters, using an iterative approach. A Probit is the same as a z-score: for example, the Probit corresponding to 70% mortality is $z_{.70}$ or = 0.52. The LC50 is calculated from the regression and is the concentration associated with 50% mortality or z = 0. To obtain a reasonably precise estimate of the LC50 with the Probit Method, the observed proportion mortalities must bracket 0.5 and the $\log_{10}$ of the tolerance should be normally distributed. To calculate the LC50 estimate and associated 95% confidence interval, two or more of the observed proportion mortalities must be between zero and one. The original percentage of mortalities should be corrected for control mortality using Abbott's formula (Section 16.2.5.4.1; Finney, 1971) before the Probit transformation is applied to the data.

16.2.5.5.2 A goodness-of-fit procedure with the Chi-square statistic is used to determine whether the data fit the Probit model. If many data sets are to be compared to one another, the Probit Method is not recommended, because it may not be appropriate for many of the data sets. This method also is only appropriate for percent mortality data sets and should not be used for estimating endpoints that are a function of the control response, such as inhibition of growth or reproduction. Most computer programs that generate Probit estimates also generate confidence interval estimates for the LC50. These confidence interval estimates on the LC50 might not be correct if replicate mortalities are pooled to obtain a mean treatment response (USEPA-USACE, 1998a). This can be avoided by entering the Probit-transformed replicate responses and doing a least-squares regression on the transformed data.

16.2.5.6 The Trimmed Spearman-Karber Method

16.2.5.6.1 The trimmed Spearman-Karber Method is a modification of the Spearman-Karber, nonparametric statistical procedure for estimating the LC50 and the associated 95% confidence interval (Hamilton et al., 1977). This procedure estimates the trimmed mean of the distribution of the $\log_{10}$ of the tolerance. If the log tolerance distribution is symmetric, this estimate of the trimmed mean is equivalent to an estimate of the median of the log tolerance distribution. Use of the trimmed Spearman-Karber Method is only appropriate for lethality data sets when the requirements for the Probit Method are not met (USEPA, 1994c; USEPA, 1994d).

16.2.5.6.2 To calculate the LC50 estimate with the trimmed Spearman-Karber Method, the smoothed, adjusted, observed proportion mortalities must bracket 0.5. To calculate a confidence interval for the LC50 estimate, one or more of the smoothed, adjusted, observed proportion mortalities must be between zero and one.

16.2.5.6.3 Smooth the observed proportion mortalities as described for the Probit Method. Adjust the smoothed observed proportion mortality in each concentration for mortality in the control group using Abbott's formula (see Probit Method, Section 16.2.5.5). Calculate the amount of trim to use in the estimation of the LC50 as follows:

$$\text{Trim} = \max(p_1^a, 1 - p_k^a)$$

where p_1^a = the smoothed, adjusted proportion mortality for the lowest treatment concentration, exclusive of the control.

p_k^a = the smoothed, adjusted proportion mortality for the highest treatment concentration.

k = the number of treatment concentrations, exclusive of the control.

16.2.5.7 Linear Interpolation Method

16.2.5.7.1 The Linear Interpolation Method calculates a toxicant concentration that causes a given percent reduction (e.g., 25%, 50%, etc.) in the endpoint of interest and is reported as an ICp value (IC = Inhibition Concentration; where p = the percent effect). The procedure was designed for general applicability in the analysis of data from chronic toxicity tests and for the generation of an endpoint from a continuous model that allows a traditional quantitative assessment of the precision of the endpoint, such as confidence limits for the endpoint of a single test or a mean and coefficient of variation for the endpoints of multiple tests.

16.2.5.7.2 As described in USEPA (1994c; 1994d), the Linear Interpolation Method of calculating an ICp assumes that the responses (1) are monotonically nonincreasing, where the mean response for each higher concentration is less than or equal to the mean response for the previous concentration, (2) follow a piecewise linear response function, and (3) are from a random, independent, and representative sample of test data. If the data are not monotonically nonincreasing, they are adjusted by smoothing (averaging). In cases where the responses at the low toxicant concentrations are much higher than in the controls, the smoothing process may result in a large upward adjustment in the control mean. In the Linear Interpolation Method, the smoothed response means are used to obtain the ICp estimate reported for the test. No assumption is made about the distribution of the data except that the data within a group being resampled are independent and identically distributed.

16.2.5.7.3 The Linear Interpolation Method assumes a linear response from one concentration to the next. Thus, the IC is estimated by linear interpolation between two concentrations whose responses bracket the response of interest, the (p) percent reduction from the control.

16.2.5.7.4 If the assumption of monotonicity of test results is met, the observed response means ($\overline{Y}_i$) should stay the same or decrease as the toxicant concentration increases. If the means do not decrease monotonically, the responses are “smoothed” by averaging (pooling) adjacent means. Observed means at each concentration are considered in order of increasing concentration, starting with the control mean ($\overline{Y}_1$). If the mean observed response at the lowest toxicant concentration ($\overline{Y}_2$) is equal to or smaller than the control mean ($\overline{Y}_1$), it is used as the response. If it is larger than the control mean, it is averaged with the control, and this average is used for both the control response (M_1) and the lowest toxicant concentration response (M_2). This mean is then compared to the mean observed response for the next higher toxicant concentration ($\overline{Y}_3$). Again, if the mean observed response for the next higher toxicant concentration is smaller than the mean of the control and the lowest toxicant concentration, it is used as the response. If it is higher than the mean of the first two, it is averaged with the mean of the first two, and the resulting mean is used as the response for the control and two lowest concentrations of toxicant. This process is continued for data from the remaining toxicant concentrations. Unusual patterns in the deviations from monotonicity may require an additional step of smoothing. Where $\overline{Y}_i$ decrease monotonically, the $\overline{Y}_i$ become M_i without smoothing.

16.2.5.7.5 To obtain the ICp estimate, determine the concentrations C_J and C_{J+1} that bracket the response M_1 (1 - p/100), where M_1 is the smoothed control mean response and p is the percent reduction in response relative to the control response. These calculations can easily be done by hand or with a computer program as described below. The linear interpolation estimate is calculated as follows:

$$ICp = C_J + [\, M_1(1 - p/100) - M_J \,] \; \frac{(C_{J+1} - C_J)}{(M_{J+1} - M_J)}$$

where C_J = tested concentration whose observed mean response is greater than M_1(1 - p/100).

C_{J+1} = tested concentration whose observed mean response is less than M_1(1 - p/100).

M_1 = smoothed mean response for the control.

M_J = smoothed mean response for concentration J.

M_{J+1} = smoothed mean response for concentration J + 1.

p = percent reduction in response relative to the control response.

IC*p* = estimated concentration at which there is a percent reduction from the smoothed mean control response.

16.2.5.7.6 Standard statistical methods for calculating confidence intervals are not applicable for the ICp. The bootstrap method, as proposed by Efron (1982), is used to obtain the 95% confidence interval for the true mean. In the bootstrap method, the test data Y_{ji} is randomly resampled with replacement to produce a new set of data Y_{ji}* that is statistically equivalent to the original data, but which produces a new and slightly different estimate of the IC*p* (IC*p**). This process is repeated at least 80 times (Marcus and Holtzman, 1988), resulting in multiple “data” sets, each with an associated ICp* estimate. The distribution of the ICp* estimates derived from the sets of resampled data approximates the sampling distribution of the ICp estimate. The standard error of the ICp is estimated by the standard deviation of the individual ICp* estimates. Empirical confidence intervals are derived from the quantiles of the ICp* empirical distribution. For ex-

ample, if the test data are resampled a minimum of 80 times, the empirical 2.5% and the 97.5% confidence limits are about the second smallest and second largest ICp* estimates (Marcus and Holtzman, 1988). The width of the confidence intervals calculated by the bootstrap method is related to the variability of the data. When confidence intervals are wide, the reliability of the IC estimate is in question. However, narrow intervals do not necessarily indicate that the estimate is highly reliable, because of undetected violations of assumptions and the fact that the confidence limits based on the empirical quantiles of a bootstrap distribution of 80 samples may be unstable.

16.2.6 Analysis of Bioaccumulation Data

16.2.6.1 In some cases, body burdens will not approach steady-state body burdens in a 28-d test (ASTM, 1999c). Organic compounds exhibiting these kinetics will probably have a log Kow >5, be metabolically refractory (e.g., highly chlorinated PCBs, dioxins), or have low depuration rates. Additionally, tissue residues of several heavy metals may gradually increase over time so that 28 d is inadequate to approach steady-state. Depending on the goals of the study and the adaptability of the test species to long-term testing, it may be necessary to conduct an exposure longer than 28 d (or a kinetic study) to obtain a sufficiently accurate estimate of steady-state tissue residues of these compounds.

16.2.6.2 Biotic Sampling

16.2.6.2.1 In the long-term studies, the exposure should continue until steady-state body burdens are attained. ASTM (1999c) recommends a minimum of five sampling periods (plus t_0) when conducting water exposures to generate bioconcentration factors (BCFs). Sampling in a geometric progression is also recommended with sampling times reasonably close to S/16, S/8, S/4, S/2, and S, where S is the time to steady state. This sampling design assumes a fairly accurate estimate of time to steady state, which is often not the case with sediment exposures.

16.2.6.2.2 To document steady state from sediment exposures, placing a greater number of samples at and beyond the predicted time to steady state is recommended. With a chemical expected to reach steady state within 28 to 50 d, samples should be taken at Day 0, 7, 14, 21, 28, 42, 56, and 70. If the time to steady state is much greater than 42 d, then additional sampling periods at two-week intervals should be added (e.g., Day 84). Slight deviations from this schedule (e.g., Day 45 versus Day 42) are not critical, though for comparative purposes, samples should be taken at t_{28}. An estimate of time to steady state may be obtained from the literature or estimated from structure-activity relationships, though these values should be considered the minimum times to steady state.

16.2.6.2.3 This schedule increases the likelihood of statistically documenting that steady state has been obtained although it does not document the initial uptake phase as well. If an accurate estimate of the sediment uptake rate coefficient (Ks) is required, additional sampling periods are necessary during the initial uptake phase (e.g., Day 0, 2, 4, 7, 10, 14).

16.2.6.3 Abiotic Samples

16.2.6.3.1 The bioavailable fraction of the contaminants as well as the nutritional quality of the sediment are more prone to depletion in extended tests than during the 28-d exposures. To statistically document whether such depletions have occurred, replicate sediment samples should be collected for physical and chemical analysis from each sediment type at the beginning and the end of the exposure. Archiving sediment samples from every biological sampling period also is recommended.

16.2.6.4 Short-term Uptake Tests

16.2.6.4.1 Compounds may attain steady state in the oligochaete, *Lumbriculus variegatus,* in less than 28 d (Kukkonen and Landrum, 1993). However, before a shorter test is used, it must be ascertained that the analytes of interest do indeed achieve steady state in *L. variegatus* in <28 d. Biotic and abiotic samples should be taken at Day 0 and 10 following the same procedure used for the 28-d tests. If time-series biotic samples are desired, sample on Day 0, 1, 3, 5, 7, and 10.

16.2.6.5 Estimating Steady State

16.2.6.5.1 In tests where steady state cannot be documented, it may be possible to estimate steady-state concentrations. Several methods have been published that can be used to predict steady-state chemical concentrations from uptake and depuration kinetics (Spacie and Hamelink, 1982; Davies and Dobbs, 1984). All of these methods were derived from fish exposures and most use a linear uptake, first-order depuration model that can be modified for uptake of chemicals from sediment. To avoid confusing uptake from water versus sediment, Ks, the sediment uptake rate coefficient, is used instead of K1. The Ks coefficient has also been referred to as the uptake clearance rate (Landrum et al., 1989). Following the recommendation of Stehly et al. (1990), the gram sediment and gram tissue units are retained in the formulation:

$$Ct\ (t) = Ks \times Cs\ /\ K2 \times (1-e^{-K2 \times t})$$

where Ct = chemical concentration in tissue at time t

Cs = chemical concentration in sediment

Ks = uptake rate coefficient in tissue (g sed g^{-1} day^{-1})

K2 = depuration constant (day^{-1})

t = time (days)

As time approaches infinity, the maximum or equilibrium chemical concentration within the organism (Ct_{max}) becomes

$$Ct_{max} = Cs \times Ks / K2$$

Correspondingly, the bioaccumulation factor (BAF) for a compound may be estimated from

$$BAF = Ks / K2$$

16.2.6.5.2 This model assumes that the sediment concentration and the kinetic coefficients are invariant. Depletion of the sediment concentrations in the vicinity of the organism would invalidate the model. Further, the rate coefficients are conditional on the environment and health of the test organisms. Thus, changes in environmental conditions such as temperature or changes in physiology such as reproduction will also invalidate the model. Despite these potential limitations, the model can provide estimates of steady-state tissue residues.

16.2.6.5.3 The kinetic approach requires an estimate of Ks and K2, which are determined from the changes in tissue residues during the uptake phase and depuration phase, respectively. The uptake experiment should be short enough that an estimate of Ks is made during the linear portion of the uptake phase to avoid an unrealistically low uptake rate due to depuration. The depuration phase should be of sufficient duration to smooth out any loss from a rapidly depurated compartment such as loss from the voiding of feces. Unless there is reason to suspect that the route of exposure will affect the depuration rate, it is acceptable to use a K2 derived from a water exposure. For further discussion of this method for bioconcentration studies in fish, see Davies and Dobbs (1984), Spacie and Hamelink (1982), and ASTM (1999b). For application of this procedure for sediment, see ASTM (1999c). Recent studies of the accumulation of sediment-associated chemicals by benthos suggest that the kinetics for freshly dosed sediments may require a more complex formulation to estimate the uptake clearance constant than that presented above (Landrum, 1989).

16.2.6.5.4 This model predicts that equilibrium would be reached only as time becomes infinite. Therefore, for practical reasons, apparent steady state is defined here as 95% of the equilibrium tissue residue. The time to reach steady state can be estimated by

$$S = \ln[1 / (1.00-0.95)] / K2 = 3.0 / K2$$

where S = time to apparent steady state (days)

Thus, the key information is the depuration rate of the compound of interest in the test species or phylogenetically related species. Unfortunately, little of this data has been generated for benthic invertebrates. When no depuration rates are available, the depuration rate constant for organic compounds can then be estimated from the relationship between Kow and K2 for fish species (Spacie and Hamelink, 1982):

$$K2 = \text{antilog}[1.47-0.414 \times \log(Kow)]$$

The relationship between S and K2 and between K2 and Kow is summarized in Table 16.2. Estimated time (days) to reach 95% of chemical steady-state tissue residue (S) and depuration rate constants (K2) are calculated from octanol-water partition coefficients using a linear uptake, first-order depuration model (Spacie and Hamelink, 1982). The K2 values are the amount depurated (decimal fraction of tissue residue lost per day). Table 16.2 may be used to make a rough estimate of the exposure time to reach steady-state tissue residues if a depuration rate constant for the compound of interest from a phylogenetically similar species is available. If no depuration rate is available, then the table may be used for estimating the S of organic compounds from the Kow value. However, as these data were developed from fish bioconcentration data, its applicability to the kinetics of uptake from sediment-associated chemicals is unknown. The portion of organics readily available for uptake may be small in comparison to the total sediment organic concentration (Landrum, 1989). Therefore S values generated by this model should be considered as minimum time periods.

16.2.6.5.5 Using a linear uptake, first-order depuration model to estimate exposure time to reach steady-state body burden for metals is problematical for a number of reasons. The kinetics of uptake may be dependent upon a small fraction of the total sediment metal load that is bioavailable (Luoma and Bryan, 1982). Depuration rates may be more difficult to determine, as metals bound to proteins may have very low exchange rates (Bryan, 1976). High exposure concentrations of some metals can lead to the induction of metal binding proteins, like metallothionein, which detoxify metals. These metal-protein complexes within the organism have extremely low exchange rates with the environment (Bryan, 1976). Thus, the induction of metal binding proteins may result in decreased depuration rate constants in organisms exposed to the most polluted sediments. Additionally, structure-activity relationships that exist for organic chemicals (e.g., relationship between Kow and BCFs) are not well developed for metals.

Table 16.2 Estimated Time to Obtain 95 Percent of Steady-state Tissue Residue

Log Kow	K2	S (days)
1	0.114	0.2
2	0.44	0.5
3	0.17	1.4
4	0.0065	3.5
5	0.0025	9.2
6	0.00097	24
7	0.00037	61
8	0.00014	160
9	0.00006	410

16.3 Data Interpretation

16.3.1 Sediments spiked with known concentrations of chemicals can be used to establish cause and effect relationships between chemicals and biological responses. Results of toxicity tests with test materials spiked into sediments at different concentrations may be reported in terms of an LC50 (median lethal concentration), an EC50 (median effect concentration), an IC50 (inhibition concentration), or as a NOEC (no observed effect concentration) or LOEC (lowest observed effect concentration; Section 3). Consistent spiking procedures should be followed in order to make interlaboratory comparisons (Section 8.3).

16.3.2 Evaluating effect concentrations for chemicals in sediment requires knowledge of factors controlling the bioavailability. Similar concentrations of a chemical in units of mass of chemical per mass of sediment dry weight often exhibit a range in toxicity in different sediments (Di Toro et al., 1991; USEPA, 1992c). Effect concentrations of chemicals in sediment have been correlated to interstitial water concentrations, and effect concentrations in interstitial water are often similar to effect concentrations in water-only exposures. The bioavailability of nonionic organic compounds are often inversely correlated with the organic carbon concentration of the sediment. Whatever the route of exposure, the correlations of effect concentrations to interstitial water concentrations indicate that predicted or measured concentrations in interstitial water can be useful for quantifying the exposure concentration to an organism. Therefore, information on partitioning of chemicals between solid and liquid phases of sediment can be useful for establishing effect concentrations.

16.3.3 Toxic units can be used to help interpret the response of organisms to multiple chemicals in sediment. A toxic unit is the concentration of a chemical divided by an effect concentration. For example, a toxic unit of exposure can be calculated by dividing the measured concentration of a chemical in pore water by the water-only LC50 for the same chemical (Ankley et al., 1991a). Toxicity expressed as toxic units may be summed and this may provide information on the toxicity of chemical mixtures (Ankley et al., 1991a).

16.3.4 Field surveys can be designed to provide either a qualitative reconnaissance of the distribution of sediment contamination or a quantitative statistical comparison of contamination among sites (Burton and Ingersoll, 1994). Surveys of sediment toxicity are usually part of more comprehensive analyses of biological, chemical, geological, and hydrographic data. Statistical correlation can be improved and costs reduced if subsamples are taken simultaneously for sediment toxicity or bioaccumulation tests, chemical analyses, and benthic community structure.

16.3.5 Descriptive methods, such as toxicity tests with field-collected sediment, should not be used alone to evaluate sediment contamination. An integration of several methods using the weight of evidence is needed to assess the effects of contaminants associated with sediment (Long and Morgan, 1990; Ingersoll et al., 1996, 1997; MacDonald et al., 1996). Hazard evaluations integrating data from laboratory exposures, chemical analyses, and benthic community assessments provide strong complementary evidence of the degree of pollution-induced degradation in aquatic communities (Chapman et al., 1992, 1997; Burton, 1991; Canfield et al., 1994, 1996, 1998).

16.3.6 Toxicity Identification Evaluation (TIE) procedures can be used to help provide insights as to specific contaminants responsible for toxicity in sediment (USEPA, 1991b; Ankley and Thomas, 1992). For example, the toxicity of contaminants such as metals, ammonia, hydrogen sulfide, and nonionic organic compounds can be identified using TIE procedures.

16.3.7 Interpretation of Comparisons of Tissue Residues

16.3.7.1 If the mean control tissue residues at Day 28 are not significantly greater than the Day 0 tissue residues, it can be concluded that there is no significant contamination from the exposure system or from the control sediment. If there is significant uptake, the exposure system or control sediment should be reevaluated as to suitability. Even if there is a significant uptake in the controls, it is still possible to compare the controls and treatments as long as the contaminant concentrations in the test tissue residues are substantially higher. However, if control values are high, the data should be discarded and the experiment conducted again after determining the source of contamination.

16.3.7.2 Comparisons of the 28-d control (or reference) tissue residues and 28-d treatment tissue residues determines whether there was statistically significant bioaccumulation due to exposure to test sediments. Comparisons between control and reference tissue residues at Day 28 determine whether there was a statistically significant bioaccumulation due to exposure to the reference sediment. If no significant difference is detected when treatment tissue residues are compared to a set criterion value (e.g., FDA action level) with a one-tailed test, the residues must be considered equivalent to the value even though numerically the mean treatment tissue residue may be smaller.

16.3.7.3 BAFs and BSAFs

16.3.7.3.1 Statistical comparisons between ratios such as BAFs or BSAFs are difficult due to computation of error terms. Since all variables used to compute BAFs and BSAFs have errors associated with them, it is necessary to estimate the variance as a function of these errors. This can be accomplished using approximation techniques such as the propagation of error (Beers, 1957) or a Taylor series expansion method (Mood et al., 1974). BAFs and BSAFs can then be compared using these estimates of the variance. ASTM (1999c) provides examples of this approach.

16.3.7.4 Comparing Tissue Residues of Different Compounds

16.3.7.4.1 In some cases, it is of interest to compare the tissue residues of different compounds. For example, Rubinstein et al. (1987) compared the uptake of thirteen different PCB congeners to test for differences in bioavailability. Because the values for the different compounds are derived from the same tissue samples, they are not independent and tend to be correlated, so standard *t* tests and ANOVAs are inappropriate. A repeated measures technique (repeated testing of the same experimental unit) should be used where the experimental unit (individual) is considered as a random factor and the different compounds as a second factor. See Rubinstein et al. (1987) and Lake et al. (1990) for an example of the application of repeated measures to bioaccumulation data.

16.4 Reporting

16.4.1 The record of the results of an acceptable sediment test should include the following information either directly or by referencing available documents:

16.4.1.1 Name of test and investigator(s), name and location of laboratory, and dates of start and end of test.

16.4.1.2 Source of control or test sediment, and method for collection, handling, shipping, storage and disposal of sediment.

16.4.1.3 Source of test material, lot number if applicable, composition (identities and concentrations of major ingredients and impurities if known), known chemical and physical properties, and the identity and concentration(s) of any solvent used.

16.4.1.4 Source and characteristics of overlying water, description of any pretreatment, and results of any demonstration of the ability of an organism to survive or grow in the water.

16.4.1.5 Source, history, and age of test organisms; source, history, and age of brood stock, culture procedures; and source and date of collection of the test organisms, scientific name, name of person who identified the organisms and the taxonomic key used, age or life stage, means and ranges of weight or length, observed diseases or unusual appearance, treatments used, and holding procedures.

16.4.1.6 Source and composition of food; concentrations of test material and other contaminants; procedure used to prepare food; and feeding methods, frequency and ration.

16.4.1.7 Description of the experimental design and test chambers, the depth and volume of sediment and overlying water in the chambers, lighting, number of test chambers and number of test organisms/treatment, date and time test starts and ends, temperature measurements, dissolved oxygen concentration (µg/L) and any aeration used before starting a test and during the conduct of a test.

16.4.1.8 Methods used for physical and chemical characterization of sediment.

16.4.1.9 Definition(s) of the effects used to calculate LC50 or EC50s, biological endpoints for tests, and a summary of general observations of other effects.

16.4.1.10 A table of the biological data for each test chamber for each treatment, including the control(s), in sufficient detail to allow independent statistical analysis.

16.4.1.11 Methods used for statistical analyses of data.

16.4.1.12 Summary of general observations on other effects or symptoms.

16.4.1.13 Anything unusual about the test, any deviation from these procedures, and any other relevant information.

16.4.2 Published reports should contain enough information to clearly identify the methodology used and the quality of the results.

Section 17
Precision and Accuracy

17.1 Determining Precision and Accuracy

17.1.1 Precision is a term that describes the degree to which data generated from replicate measurements differ and reflects the closeness of agreement between randomly selected test results. Accuracy is the difference between the value of the measured data and the true value and is the closeness of agreement between an observed value and an accepted reference value. Quantitative determination of precision and accuracy in sediment testing of aquatic organisms is difficult or may be impossible in some cases, as compared to analytical (chemical) determinations. This is due, in part, to the many unknown variables that affect organism response. Determining the accuracy of a sediment test using field samples is not possible since the true values are not known. Since there is no acceptable reference material suitable for determining the accuracy of sediment tests, the accuracy of the test methods has not been determined (Section 17.2).

17.1.2 Sediment tests exhibit variability due to several factors (Section 9). Test variability can be described in terms of two types of precision, either single laboratory (intralaboratory or repeatability; Section 17.5.1) precision or multi-laboratory (interlaboratory or reproducibility; Section 17.5.2, 17.5.3 and 17.6) precision. Intralaboratory precision reflects the ability of trained laboratory personnel to obtain consistent results repeatedly when performing the same test on the same organism using the same toxicant. Interlaboratory precision (also referred to as round-robin or ring tests) is a measure of the reproducibility of a method when tests are conducted by a number of laboratories using that method and the same organism and samples. Generally, intralaboratory results are less variable than interlaboratory results (USEPA, 1991a; USEPA, 1991c; USEPA, 1994b; USEPA, 1994c; Hall et al., 1989; Grothe and Kimerle, 1985).

17.1.3 A measure of precision can be calculated using the mean and relative standard deviation (percent coefficient of variation, or CV% = standard deviation/mean x 100) of the calculated endpoints from the replicated endpoints of a test. However, precision reported as the CV should not be the only approach used for evaluating precision of tests and should not be used for the NOEC levels derived from statistical analyses of hypothesis testing. The CVs can be very high when testing extremely toxic samples. For example, if there are multiple replicates with no survival and one with low survival, the CV might exceed 100%, yet the range of response is actually quite consistent. Therefore, additional estimates of precision should be used, such as range of responses, and minimum detectable differences (MDD) compared to control survival or growth. Several factors can affect the precision of the test, including test organism age, condition and sensitivity; handling and feeding of the test organisms; overlying water quality; and the experience of the investigators in conducting tests. For these reasons, it is recommended that trained laboratory personnel conduct the tests in accordance with the procedures outlined in Section 9. Quality assurance practices should include the following: (1) single laboratory precision determinations that are used to evaluate the ability of the laboratory personnel to obtain precise results using reference toxicants for each of the test organisms and (2) preparation of control charts (Section 17.4) for each reference toxicant and test organism. The single laboratory precision determinations should be made before conducting a sediment test and should be periodically performed as long as whole-sediment tests are being conducted at the laboratory.

17.1.4 Intralaboratory precision data are routinely calculated for test organisms using water-only 96-h exposures to a reference toxicant, such as potassium chloride (KCl). Intralaboratory precision data should be tracked using a control chart. Each laboratory's reference-toxicity data will reflect conditions unique to that facility, including dilution water, culturing, and other variables (Section 9). However, each laboratory's reference-toxicity CVs should reflect good repeatability.

17.1.5 Interlaboratory precision (round-robin) tests have been completed with both *Hyalella azteca* and *Chironomus tentans* using 4-d water-only tests and 10-d whole-sediment tests described in Section 11.2 and 12.2 (Section 17.5). Section 17.6 describes results of round-robin evaluations with long-term sediment toxicity tests described in Sections 14 and 15 for *H. azteca* and *C. tentans.*

17.2 Accuracy

17.2.1 The relative accuracy of toxicity tests cannot be determined since there is no acceptable reference material. The relative accuracy of the reference-toxicity tests can only be evaluated by comparing test responses to control charts.

17.3 Replication and Test Sensitivity

17.3.1 The sensitivity of sediment tests will depend in part on the number of replicates per concentration, the probability levels (alpha and beta) selected, and the type of statistical analysis. For a specific level of variability, the sensitivity of the test will increase as the number of replicates is increased. The minimum recommended number of replicates varies with the objectives of the test and the statistical method used for analysis of the data (Section 16).

17.4 Demonstrating Acceptable Laboratory Performance

17.4.1 Intralaboratory precision, expressed as a coefficient of variation (CV), can be determined by performing five or more tests with different batches of test organisms, using the same reference toxicant, at the same concentrations, with the same test conditions (e.g., the same test duration, type of water, age of test organisms, feeding), and same data analysis methods. A reference-toxicity concentration series (dilution factor of 0.5 or higher) should be selected that will provide partial mortalities at two or more concentrations of the test chemical (Section 9.14, Table 9.1, 9.2). See Section 9.16 for additional detail on reference-toxicity testing.

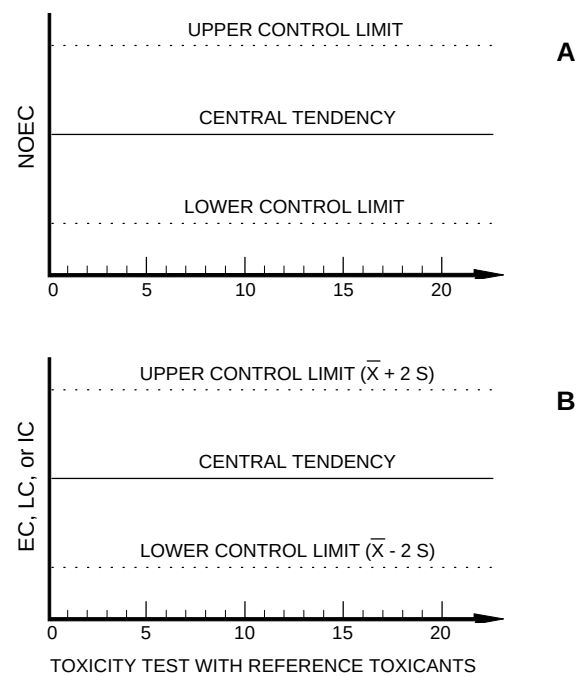
17.4.2 It is desirable to determine the sensitivity of test organisms obtained from an outside source. The supplier should provide data with the shipment describing the history of the sensitivity of organisms from the same source culture.

17.4.3 Before conducting tests with potentially contaminated sediment, it is strongly recommended that the laboratory conduct the tests with control sediment(s) alone. Results of these preliminary studies should be used to determine if use of the control sediment and other test conditions (i.e., water quality) result in acceptable performance in the tests as outlined in Tables 11.1, 12.1, 13.1, 14.1, and 15.1.

17.4.4 A control chart should be prepared for each combination of reference toxicant and test organism. Each control chart should include the most current data. Endpoints from five tests are adequate for establishing the control charts. In this technique, a running plot is maintained for the values (X_i) from successive tests with a given reference toxicant (Figure 17.1), and the endpoints (LC50, NOEC, ICp) are examined to determine if they are within prescribed limits. Control charts as described in USEPA (1991a) and USEPA (1993b) are used to evaluate the cumulative trend of results from a series of samples. The mean and upper and lower control limits (±2 SD) are recalculated with each successive test result. After two years of data collection, or a minimum of 20 data points, the control (cusum) chart should be maintained using only the 20 most recent data points.

17.4.5 The outliers, which are values falling outside the upper and lower control limits, and trends of increasing or decreasing sensitivity, are readily identified using control charts. With an alpha of 0.05, one in 20 tests would be expected to fall outside of the control limits by chance alone. During a 30-d period, if two reference-toxicity tests out of a total of the previous 20 fall outside the control limits, the sediment toxicity tests conducted during the time in which the second reference-toxicity test failed are suspect and should be considered as provisional and subject to careful review.

17.4.5.1 A sediment test may be acceptable if specified conditions of a reference-toxicity test fall outside the expected ranges (Section 9). Specifically, a sediment test should not necessarily be judged unacceptable if the LC50 for a given reference-toxicity test falls outside the expected range or if mortality in the control of the reference-toxicity test exceeds 10% (Tables 9.1 and 9.2). All the performance criteria outlined in Tables 11.3, 12.3, 13.4,



$$\bar{x} = \frac{\sum_{i=1} x_i}{n}$$

$$s = \sqrt{\frac{\sum_{i=1}^{n} x_i^2 - \frac{(\sum_{i=1}^{n} x_i)^2}{n}}{n-1}}$$

where x_1 = Successive toxicity values of toxicity tests.
n = Number of tests.
$\bar{x}$ = Mean toxicity value.
S = Standard deviation.

Figure 17.1 Control (cusum) charts: (A) hypothesis testing results; and (B) point estimates (LC, EC, or IC).

14.3, and 15.3 must be considered when determining the acceptability of a sediment test. The acceptability of the sediment test would depend on the experience and judgment of the investigator and the regulatory authority.

17.4.6 If the value from a given test with the reference toxicant falls more than two standard deviations (SD) outside the expected range, the sensitivity of the organisms and the overall credibility of the test system may be suspect (USEPA, 1991a). In this case, the test procedure should be examined for defects and should be repeated with a different batch of test organisms.

17.4.7 Performance should improve with experience, and the control limits for point estimates should gradually narrow. However, control limits of ±2 SD, by definition, will be exceeded 5% of the time, regardless of how well a laboratory performs. Highly proficient laboratories that develop a very narrow control limit may be unfairly penalized if a test that falls just outside the control limits is rejected *de facto*. For this reason, the width of the control limits should be considered in determining whether or not an outlier is to be rejected. This determination may be made by the regulatory authority evaluating the data.

17.4.8 The recommended reference-toxicity test consists of a control and five or more concentrations in which the endpoint is an estimate of the toxicant concentration that is lethal to 50% of the test organisms in the time period prescribed by the test. The LC50 is determined by an appropriate procedure, such as the trimmed Spearman-Karber Method, Probit Method, Graphical Method, or the Linear Interpolation Method (Section 16).

17.4.9 The point estimation analysis methods recommended in this manual have been chosen primarily because they are well-tested, well-documented, and are applicable to most types of test data. Many other methods were considered in the selection process, and it is recognized that the methods selected are not the only possible methods of analysis of toxicity data.

17.5 Precision of Sediment Toxicity Test Methods: Evaluation of 10-d Sediment Tests and Reference-toxicity Tests

17.5.1 Intralaboratory Performance

17.5.1.1 Intralaboratory performance of the *Hyalella azteca* and *Chironomus tentans* 10-d tests (as described in Tables 11.1 and 12.1) was evaluated at the USEPA Office of Research and Development Laboratory (Duluth, MN) using one control sediment sample in June 1993. In this study, five individuals simultaneously conducted the 10-d whole-sediment toxicity tests as described in Tables 11.1 and 12.1 with the exception of the feeding rate of 1.0 mL rather than 1.5 mL for *C. tentans*. The results of the study are presented in Table 17.1. The mean survival for *H. azteca* was 90.4% with a CV of 7.2% and the mean survival for *C. tentans* was 93.0% with a CV of 5.7%. All of the individuals met the survival performance criteria of 80% for *H. azteca* (Table 11.3) or 70% for *C. tentans* (Table 12.3).

17.5.2 Interlaboratory Precision: 1993 Evaluation of the 10-d Sediment Tests and the Reference-toxicity Tests

17.5.2.1 Interlaboratory precision using reference-toxicity tests or 10-d whole-sediment toxicity tests using the methods described in this manual (Tables 9.1, 9.2, 11.1, and 12.1) were conducted by federal government laboratories, contract laboratories, and academic laboratories that had demonstrated experience in sediment toxicity testing for a first time in 1993 (Section 17.5.2.2 and Burton et al., 1996b) and a second time in 1996/1997 (the "1996/1997 study"; Section 17.5.3). In the 1993 study the only exception to the methods outlined in Table 9.1 and 9.2 was that 80% rather than the current recommendation of 90% survival was used to judge the acceptability of the reference-toxicity tests. The 1993 round-robin study was conducted in two phases for each test organism. The experimental design for the 1993 round-robin study required each laboratory to conduct 96-h water-only reference-toxicity tests in Phase 1 and 10-d whole-sediment tests in Phase 2 with *Hyalella azteca* or *Chironomus tentans* over a period of six months. Criteria for selection of participants in the 1993 round-robin study were that the laboratories: (1) had existing cultures of the test organisms, (2) had experience conducting tests with the organisms, and (3) would participate voluntarily. The test methods for the reference-toxicity tests and the whole-sediment toxicity tests were similar among laboratories. Standard operating procedures detailing the test methods were provided to all participants. Culture methods were not specified and were not identical across laboratories.

Table 17.1 Intralaboratory Precision for Survival of *Hyalella azteca* and *Chironomus tentans* in 10-d Whole-sediment Toxicity Tests, June 1993[1]

	Percent Survival	
Individual	*H. azteca*	*C. tentans*
A	85	85
B	93	93
C	90	93
D	84	94
E	100	100
N	5	5
Mean	90.4	93.0
CV	7.2%	5.7%

[1] Test sample was from a control sediment (T.J. Norberg-King, USEPA, Duluth, MN, personal communication). The test was conducted at the same time by five individuals at the USEPA Office of Research and Development Laboratory (Duluth, MN). The source of overlying water was from Lake Superior.

Table 17.2 Participants in 1993 Round-robin Studies[1]

	Chironomus tentans			*Hyalella azteca*	
Laboratory	96-h KCl Test Dec 92	96-h KCl Test May 93	10-d Sediment Test May 93	96-h KCl Test Oct 92	10-d Sediment Test Mar 93
A	Y	N	N	Y	N
B	Y	Y	Y	Y	Y
C	Y	N	Y	Y	Y
D	Y	Y	Y	N	N
E	Y	Y	Y	Y	Y
F	Y	Y	Y	Y	Y
G	Y	Y	Y	Y	Y
H	Y	N	N	Y	N
I	Y	Y	Y	—[2]	Y
J	Y	Y	Y	Y	Y
K	—[3]	—[3]	—[3]	Y	Y
L	—[4]	—[4]	—[4]	Y	Y
N	10	7	8	10	9

[1] Y = Laboratory participated in testing sediment samples.
[2] Test in January 1993.
[3] Participated using *C. riparius* only.
[4] Did not intend to participate with *C. tentans.*

Table 17.3 Interlaboratory Precision for *Hyalella azteca* 96-h LC50s from Water-only Static Acute Toxicity Tests Using a Reference Toxicant (KCl) (October 1992)

Laboratory	KCl LC50 (mg/L)	Confidence Intervals Lower	Confidence Intervals Upper	Percent Control Survival
A	372	352	395	100
B	321	294	350	98
C	232	205	262	100
D	—[1]	—[1]	—[1]	—[1]
E	325	282	374	100
F	276	240	316	98
G	297	267	331	73
H	336	317	356	100
I	142[2]	101	200	93
J	337	286	398	100
L	250	222	282	100
N	10			10
Mean 1	289.0[3]			96.2%
CV 1	23.0%[3]			8.3%
N	9			9
Mean 2	305.0[4]			98.8
CV 2	14.2%[4]			2.1%

[1] Laboratory did not participate in *H. azteca* test in October.
[2] Results are from a retest in January using three concentrations only; results excluded from analysis.
[3] Mean 1 and CV 1 include all data points
[4] Mean 2 and CV 2 exclude data points for all sediment samples from laboratories that did not meet minimum control survival of ≥80%.

17.5.2.2 In the second series of round-robin tests conducted in 1996/1997, 10-d and long-term toxicity testing methods were evaluated with *Hyalella azteca* and *Chironomus tentans*. Results from these interlaboratory comparisons conducted in 1996/1997 are presented in detail in Sections 17.5.3 and 17.6. The second series of interlaboratory comparisons conducted in 1996/1997 did not restrict testing to laboratories with experience. As in 1993, the participants in the 1996/1997 round-robin study included government, contract, and academic laboratories. In the 1996/1997 study, no water-only reference-toxicity tests were conducted.

17.5.2.3 Ten laboratories participated in the *H. azteca* reference-toxicity test in the 1993 study (Table 17.2). The results from the tests with KCl are summarized in Table 17.3. The test performance criteria of ≥80% control survival was met by 90% of the laboratories resulting in a mean control survival of 98.8% (CV = 2.1%). The mean LC50 was 305 mg/L (CV = 14.2%) and the LC50s ranged from 232 to 372 mg/L KCl.

17.5.2.4 In the 10-d whole-sediment tests with *H. azteca*, nine laboratories tested the three sediments described above and five laboratories tested a fourth sediment from a heavily contaminated site in the 1993 study (Table 17.4). All laboratories completed the tests; however, Laboratory C had 75% survival, which was below the acceptable test criteria for survival (Table 11.3). For these tests, the CV was calculated using the mean percent survival for the eight laboratories that met the performance criteria for the test. The CV for survival in the control sediment (RR 3) was 5.8% with a mean survival of 94.5% and survival ranging from 86% to 100%. For sediments RR 2 and RR 4, the mean survival was 3.3% and 4.3%, respectively (Table 17.4). For RR 2, survival ranged from 0% to 24% (CV = 253%) and for RR 4, the survival ranged from 0% to 11% (CV = 114%). Survival in the moderately contaminated sediment (RR 1) was 54.2% with survival ranging from 23% to 76% (CV = 38.9%). When the RR 1 data for each laboratory were compared to the control for that laboratory, the range for the minimum detectable difference (MDD) between the test sediments and the control sediment ranged from 5 to 24% with a mean of 11% (SD = 6).

17.5.2.5 The Phase 1 *C. tentans* reference-toxicity test was conducted with KCl on two occasions in the 1993 study (Tables 17.5 and 17.6). Both tests were conducted in 20 mL of test solution in 30-mL beakers using 10 replicates per treatment with 1 organism per beaker. Animals were fed 0.25 mL of a 4 g/L solution of Tetrafin® on Day 0 and Day 2 (Table 9.1). For the first reference-toxicity test comparison, 10 laboratories participated, and

Table 17.4 Interlaboratory Precision for Survival of *Hyalella azteca* in 10-d Whole-sediment Toxicity Tests Using Four Sediments (March 1993)

	Mean Percent Survival (SD) in Sediment Samples							
Laboratory	RR 1		RR 2		RR 3 (Control)		RR 4	
A	—[1]		—[1]		—[1]		—[1]	
B	76.2	(20.7)	2.5	(7.1)	97.5	(4.6)	11.2	(13.6)
C	57.52[2]	(14.9)	1.2[2]	(0)	75.0[2]	(17.7)	1.2[2]	(0)
D	—[1]		—[1]		—[1]		—[1]	
E	46.2	(17.7)	0	(0)	97.5	(7.1)	—	
F	72.5	(12.8)	23.7	(18.5)	98.7	(3.5)	0	(0)
G	50.0	(28.3)	0	(0)	100	(0)	3.3	(5.2)
H	—[1]		—[1]		—[1]		—[1]	
I	73.7	(32.0)	0	(0)	86.2	(10.6)	—	
J	65.0	(9.3)	0	(0)	96.2	(5.2)	2.5	(7.1)
K	22.5	(18.3)	0	(0)	95.0	(5.3)	—	
L	27.5	(16.7)	0	(0)	86.2	(18.5)	—	
N	9		9		9		5	
Mean 1[3]	54.6		3.0		93.0		3.6	
CV 1	36.2%		256%		9.0%		121%	
N	8		8		8		4	
Mean 2[4]	54.2		3.3		94.5		4.3	
CV 2	38.9%		253%		5.8%		114%	

[1] Laboratory did not participate in *H. azteca* test in March.
[2] Survival in control sediment (RR 3) below minimum acceptable level.
[3] Mean 1 and CV 1 include all data points.
[4] Mean 2 and CV 2 exclude data points for all sediment samples from laboratories that did not meet minimum control survival of ≥80%.

eight laboratories met the survival criteria of the round robin, which was ≥80% survival (Table 17.5). The mean LC50 for the eight laboratories that met the survival criterion was 4.25 g/L (CV of 51.8%). The LC50s ranged from 1.25 to 6.83 g/L. Length and instar were determined for a subset of organisms at the start of the tests for some of the laboratories. When length was correlated with the LC50, the larger animals were less sensitive than the smaller animals. The effect level was significantly correlated ($r^2 = 0.78$) with the organism size, which ranged from 1.56 mm to 10.87 mm (ages of animals ranged from 7- to 13-d post-deposition). The majority of these animals were the third instar, with the smallest animals in their first instar and the largest animals a mix of third and fourth instar (Table 17.5) as determined by head capsule width.

17.5.2.6 For the second Phase 1 KCl reference-toxicity tests with *C. tentans,* seven laboratories participated in the 1993 study (Table 17.6). The test conditions were identical to those in the previous reference-toxicity test except that a minimum size was specified rather than using initial age of the animals. Each laboratory was instructed to start the test when larvae were at least 0.4 to 0.6 mm long. Therefore, a more consistent size of test organisms was used in this test. Six out of the seven laboratories met the ≥80% control survival criterion with a mean LC50 of 5.37 g/L (CV = 19.6%). The LC50s ranged from 3.61 to 6.65 g/L.

17.5.2.7 Eight laboratories participated in the 10-d whole-sediment testing with *C. tentans*. The same three sediments used in the *H. azteca* whole-sediment test were used for this test in the 1993 study (Table 17.7). All test conditions were those as described in Table 12.1 with the exception of the feeding rate of 1.0 mL rather than 1.5 mL for *C. tentans.* Three laboratories did not meet the control criteria for acceptable tests of ≥70% survival in the control (RR 3) sediment (Table 12.3). For the five laboratories that successfully completed the tests, the mean survival in the control sediment (RR 3) was 92.0% (CV of 8.3%) and survival ranged from 81.2% to 98.8%. For the RR 2 sediment sample, the mean survival among the five laboratories was 3.0% (CV = 181%) and for the RR 1 sediment sample, the mean survival was 86.8% (CV = 13.5%). A significant effect on survival was not evident for the RR 1 sample, but growth was affected (Table 17.8). When the RR 1 data for each laboratory were compared to the control for that laboratory, the MDD for survival among laboratories ranged from 2.3 to 12.1% with a mean of 8% (SD = 4).

17.5.2.8 For *C. tentans,* growth in 10-d tests is a sensitive indicator of sediment toxicity (Ankley et al., 1993) and growth was also measured in the round-robin comparison in the 1993 study (Table 17.8). Using the data from five laboratories with acceptable control survival in the control sediment (RR 3), the mean weight of *C. tentans* for the control sediment (RR 3) was 1.254 mg (CV

Table 17.5 Interlaboratory Precision for *Chironomus tentans* 96-h LC50s from Water-only Static Acute Toxicity Tests Using a Reference Toxicant (KCl) (December 1992)

Laboratory	KCl LC50 (g/L)	Confidence Interval Lower	Confidence Interval Upper	Control Survival (%)	Mean Length (mm)	Instar at Start of Test	Age at Start of Test (day)
A	6.19	5.37	7.13	75[1]	10.87	3,4	1
B	6.83	6.38	7.31	100	10.43	3	13
C	5.00	4.16	6.01	100	5.78	3	11
D	3.17	2.29	4.40	100	5.86	3	11
E	2.00[2]	—[2]	—	80	6.07	3	11
F	1.25	—[3]	—	80	1.56	1	12
G	6.28	5.26	7.50	95	7.84	3	11
H	2.89	2.39	3.50	95	6.07	3	7
I	6.66	6.01	7.24	100	—[4]	—[4]	10
J	1.77	0.59	5.26	65[1]	4.42	2,3	7
N	10			10	8		10
Mean 1[5]	4.20			89.0	6.6		10.3
CV 1	52.7%			14.5%	46.6%		17.9%
N	8			8	7		8
Mean 2[6]	4.25			93.8	6.2		10.75
CV 2	51.8%			9.3%	39.5%		15.2%

[1] Control survival below minimum acceptable level.
[2] Unable to calculate LC50 with trimmed Spearman Karber; no confidence interval could be calculated.
[3] Confidence intervals cannot be calculated as no partial mortalities occurred.
[4] No animals were measured.
[5] Mean 1 and CV 1 include all data points.
[6] Mean 2 and CV 2 exclude data points for all samples from laboratories that did not meet minimum control survival of ≥80%.

Table 17.6 Interlaboratory Precision for *Chironomus tentans* 96-h LC50s from Water-only Static Acute Toxicity Tests Using a Reference Toxicant (KCl) (May 1993)

Laboratory	KCl LC50 (g/L)	Confidence Interval Lower	Confidence Interval Upper	Control Survival (%)	Age at Start of Test (day)
A	—[1]	—	—	—	—
B	6.65	—[2]	—	90	12
C	—[1]	—	—	—	—
D	5.30	4.33	6.50	55[3]	10
E	5.11	4.18	6.24	100	11
F	3.61	2.95	4.42	90	10
G	5.36	4.43	6.49	93	12
H	—[1]	—	—	—	
I	5.30	4.33	6.52	95	10-11
J	6.20	4.80	7.89	100	13
N	7			7	7
Mean 1[4]	5.36			89	11.1
CV 1	17.9%			17.5%	9.46%
N	6			6	6
Mean 2[5]	5.37			94.7	11.2
CV 2	19.6%			4.8%	9.13%

[1] Did not participate in reference-toxicity test in April.
[2] Confidence intervals cannot be calculated as no partial mortalities occurred.
[3] Control survival below minimum acceptable level.
[4] Mean 1 and CV 1 include all data points.
[5] Mean 2 and CV 2 exclude data points for all samples from laboratories that did not meet minimum control survival of ≥70%.

Table 17.7 Interlaboratory Precision for Survival of *Chironomus tentans* in 10-d Whole-sediment Toxicity Tests Using Three Sediments (May 1993)

	Mean Percent Survival (SD) in Sediment Samples					
Laboratory	RR 1		RR 2		RR 3 (Control)	
A	—[1]		—[1]		—[1]	
B	67.5	(14.9)	2.5	(7.1)	98.8	(3.5)
C	15.0[2]	(12.0)	0[2]	(0)	62.5[2]	(26.0)
D	60.0[2]	(20.0)	0[2]	(0)	66.3[2]	(27.7)
E	85.0	(11.9)	0	(0)	93.8	(9.2)
F	87.5[2]	(12.5)	0[2]	(0)	43.8[2]	(30.2)
G	90.0	(13.1)	12.5	(3.5)	87.5	(10.3)
H	—[1]		—[1]		—[1]	
I	97.5	(4.6)	0	(0)	98.8	(3.5)
J	93.8	(11.8)	0	(0)	81.2	(8.3)
N	8		8		8	
Mean 1[3]	74.5		1.88		79.1	
CV 1	36.7%		233%		25.1%	
N	5		5		5	
Mean 2[4]	86.8		3.0		92.0	
CV 2	13 5%		181%		8.3%	

[1] Did not participate in *C. tentans* test in May.
[2] Survival in control sediment (RR 3) below minimum acceptable level.
[3] Mean 1 and CV 1 include all data points.
[4] Mean 2 and CV 2 exclude data points for all sediment samples from laboratories that did not meet minimum control survival of ≥70%.

Table 17.8 Interlaboratory Precision for Growth of *Chironomus tentans* in 10-d Whole-sediment Toxicity Tests Using Three Sediments (May 1993)

	Growth—Dry Weight in mg (SD) in Sediment Samples					
Laboratory	RR 1		RR 2		RR 3 (Control)	
A	—[1]		—[1]		—[1]	
B	0.370	(0.090)	0	(0)	1.300	(0.060)
C	0.883[2]	(0.890)	0[2]	(0)	0.504[2]	(0.212)
D	0.215[2]	(0.052)	0[2]	(0)	1.070[2]	(0.107)
E	0.657	(0.198)	0	(0)	0.778	(0.169)
F	0.210[2]	(0.120)	0[2]	(0)	0.610[2]	(0.390)
G	0.718	(0.114)	0	(0)	1.710	(0.250)
H	—[1]		—[1]		—[1]	
I	0.639	(0.149)	0	(0)	1.300	(0.006)
J	0.347	(0.050)	0	(0)	1.180	(0.123)
N	8		8		8	
Mean 1[3]	0.505		—		1.056	
CV 1	49.9%		—		38.3%	
N	5		5		5	
Mean 2[4]	0.546		—		1.254	
CV 2	31.9%		—		26.6%	

[1] Did not participate in testing in May.
[2] Survival in control sediment (RR 3) below minimum acceptable level.
[3] Mean 1 and CV 1 include all data points.
[4] Mean 2 and CV 2 exclude data points for all sediment samples from laboratories that did not meet minimum control survival of ≥70%.

= 26.6%). The *C. tentans* in the moderately contaminated sediment (RR 1) had a mean weight of 0.546 mg (CV = 31.9%). No growth measurements were obtained for *C. tentans* in sediment RR 2 because of the high mortality. The mean minimum detectable difference for growth among laboratories meeting the survival performance criteria was 11% (SD = 5) and the MDD ranged from 4.8 to 23.6% when the RR 1 data were compared to the RR 3 data.

17.5.3 Interlaboratory Precision: 1996/1997 Evaluation of 10-d Sediment Tests

17.5.3.1 ***The 1996/1997 Precision Evaluation: 10-d Whole-sediment Toxicity Testing***. The results of the 10-d toxicity interlaboratory comparisons conducted in 1996/1997 are presented in Tables 17.9 to 17.12. A total of 18 laboratories participated in the 1996/1997 study; however, not all samples were tested by all laboratories. Laboratories performed the tests during a specified time period and followed methods outlined in Tables 11.1 and 12.1. Field samples were pretested to identify moderately toxic samples. Samples were prepared and subsampled at one time to increase consistency among the subsamples. Samples were shipped to the testing laboratories by express mail. Laboratories used their own water supplies and were asked to use moderately hard water (hardness about 100 mg/L as $CaCO_3$). The following samples were evaluated in the 10-d toxicity tests: a field control sediment from West Bearskin Lake, MN (WB), a formulated sediment (FS, formulated with alpha-cellulose; Kemble et al., 1999), two contaminated sediments (Little Scioto River, OH (LS); Defoe Creek site, Keweenaw, MI (DC)), and FS spiked with three concentrations of cadmium (0.3, 1.0, and 3.0 mg/kg Cd). The LS sample was primarily contaminated with polycyclic aromatic hydrocarbons and the DC sample was primarily contaminated with copper. Some laboratories did not conduct tests on all samples due to logistical constraints. In addition, ash-free dry weight (AFDW) was not measured by laboratories which did not have access to a muffle furnace.

17.5.3.2 ***The 1996/1997 Precision Evaluation – Hyalella azteca***. Eighteen laboratories participated in the 1996/1997 *H. azteca* 10-d comparison (Table 17.9). A total of 82% of the laboratories had acceptable survival (≥80%) and for these tests the average survival (and CV) was 92% (CV=5%) in the WB control sediment and 89% (CV=12%) in the formulated sediment (FS). The two contaminated field sediments (DC, LS) were moderately toxic, with the mean survival of 45% (CV=38%) in DC sediment and 57% (CV=49%) in LS sediment. The mean MDDs of the two contaminated samples for all laboratories relative to the WB control sediment were low (14% for both the DC and the LS sediments). The range of MDDs relative to the WB control sediment among all laboratories was 8 to 23% for the DC sediment and 2 to 22% for the LS sediment. A dose response effect was observed with the Cd-spiked formulated sediments. Moderate toxicity was observed in the 1 mg/kg Cd sample with a mean survival of 49% (CV=40%). The mean MDD and range for the 1 mg/kg Cd sample for all laboratories was 16% (5.7 to 26%). It is apparent from the MDDs that some laboratories had low variability while others had only moderate levels of variability.

17.5.3.3 ***The 1996/1997 Precision Evaluation – Chironomus tentans***. Eighteen laboratories participated in the 1996/1997 *C. tentans* 10-d survival and growth comparison (Table 17.10) with the same samples used in the toxicity test as described above. A total of 15 laboratories (89%) had acceptable survival (≥70%), and for these tests, the mean survival was 89% (CV=9.4%) in the WB control sediment and 88% (CV=10.2%) in the formulated sediment (FS). The two contaminated field sediments were only slightly toxic to the midge (mean survival of 80% (CV=16%) for the DC sediment and 71% (CV=33%) for LS sediment). The mean MDDs relative to the WB control sediment, across all laboratories for the two contaminated samples were low (12% for the DC sediment and 11% for LC sediment). The range of MDDs relative to the WB control sediment among laboratories were 6.1 to 22% for the DC sediment and 5.1 to 18% for LS sediment. No toxicity was observed for survival in the cadmium tests. The mean survival of midge in the 1 mg/kg Cd treatment was 92% (CV=5.6%). The mean MDD and range for the 1 mg/kg Cd sample was 12% (6.9 to 30%). It is apparent from the MDDs that some laboratories had low variability while others had slightly lower variability.

17.5.3.4 Growth of *C. tentans* was evaluated by up to 16 laboratories in 1996/1997, depending on the sample and whether or not they had capabilities to determine AFDW. For dry weight analyses, 12 of 15 laboratories had acceptable dry weight (≥0.6 mg/individual) and survival >70% in the WB control sediment, while 12 of 15 of the laboratories had acceptable dry weight and survival in the formulated sediment (FS; Table 17.11). For AFDW, 7 of 11 laboratories had acceptable weight (≥0.48 mg/individual) and survival >70% in WB control sediment (field control) (WB) and 7 of 11 laboratories reported acceptable weight in the formulated sediment (FS; Table 17.12). For the midges, the mean dry weight was 1.39 mg/organism (CV=33%) in the WB control sediment and 1.50 mg/organism (CV=31%) in the formulated sediment (FS) for laboratories that met the control survival in WB control sediment. For AFDW, mean AFDW was 0.92 mg/organism (CV=30%) in the WB control sediment and 1.161 mg/organism (CV=33%) in the formulated sediment (FS). Exposure to the contaminated DC sediment reduced the weight of the midge (mean weight of 0.49 mg/organism (CV=60%) as dry weight, while the mean weight of 0.24 mg/organism (CV=45%) was determined for the AFDW), yet exposure to LS sediment did not reduce weight of midges (1.45 mg dry weight (CV=45%); 0.86 mg AFDW (CV=27%)). The mean MDDs relative to WB control sediment, across all laboratories for the two contaminated samples, were low (0.17 mg/organism dry weight for the DC sediment and 0.28 mg dry weight for LS sediment). The range of MDDs among laboratories for dry weight was 0.04 to 0.53 mg/organism for DC sediment and 0.09 to 1.04 mg/organism for LS sediment. The AFDW data exhibited a similar pattern. Mean MDD as AFDW was 0.12 mg for the DC sediment and 0.16 mg for the LS sediment. The range

Table 17.9 Interlaboratory Precision for Survival (%) of *Hyalella azteca* in 10-d Whole-sediment Toxicity Tests (1996/1997)

	Mean Percent Survival (SD) in Sediment Samples and Cd-spiked Control Sediment													
	Sediment								Cadmium -FS Spikes (mg/kg)					
Laboratory	WB		DC		LS		FS		0.3-Cd		1-Cd		3-Cd	
A	71[a]	(23.0)	0[a]		NT[b]		40[a]	(37.8)	NT		NT		NT	
B	75[a]	(24.5)	49[a]	(27.5)	84[a]	(30.7)	90[a]	(7.6)	NT		NT		NT	
C	NT	NT	NT		NT		95[c]	(5.8)	90[c]	(14.1)	73[c]	(9.6)	0[c]	
E	85	(15.1)	31	(19.6)	71	(34.4)	83	(14.9)	68	(9.6)	83	(9.6)	3	(5.0)
F	94	(5.2)	31	(18.1)	19	(16.4)	60	(20.0)	40	(8.2)	28	(5.0)	3	(5.0)
G	83	(15.8)	38	(15.8)	28	(12.8)	90	(9.3)	NT		NT		NT	
H	95	(7.6)	61	(19.6)	64	(20.7)	99	(3.5)	NT		NT		NT	
I	95	(5.4)	33	(13.8)	85	(9.3)	99	(3.5)	83	(20.6)	28	(17.1)	0	
K	95	(7.6)	79	(9.9)	94	(7.4)	100	(0)	98	(5.0)	60	(8.2)	0	
M	86	(17.7)	23	(21.9)	50	(22.7)	85	(16.9)	80	(14.1)	65	(19.2)	0	
N	91	(6.4)	48	(10.4)	29	(23.6)	85	(14.1)	100	(0)	70	(8.2)	3	(5.0)
O	91	(8.4)	50	(14.1)	74	(10.6)	95	(5.4)	78	(22.2)	55	(26.5)	0	
P	88	(7.1)	56	(27.2)	60	(27.3)	85	(10.7)	83	(16.2)	48	(16.2)	0	
Q	91	(8.4)	20	(16.0)	84	(22.0)	96	(5.2)	98	(5.0)	23	(28.7)	0	
S	68[a]	(17.5)	34[a]	(24.5)	80[a]	(23.9)	70[a]	(25.1)	NT		NT		NT	
U	94	(7.4)	60	(30.2)	63	(21.2)	95	(5.4)	88	(12.6)	38	(15.0)	0	
V	95	(10.0)	35	(20.8)	75	(20.8)	93	(15.0)	93	(5.0)	40	(14.1)	0	
X	99	(3.5)	59	(12.5)	0		85	(15.1)	NT		NT		NT	
N-1[d]	17		17		16		17		11		11		11	
Mean-1	88		42		60		85		83		49		1	
SD-1	9.1		18.9		27.4		15.7		17.2		19.4		1.4	
CV-1	10.3		45.6		45.7		18.4		20.9		39.7		171.3	
N-2[e]	14		14		14		14		11		11		11	
Mean- 2	92		45		57		89		83		49		1	
SD- 2	4.6		17.1		27.9		10.4		17.2		19.4		1.4	
CV-2	5.0		38.3		49.1		11.6		20.9		39.7		171.3	

[a] Control survival below acceptable level of 80% in WB sediment.

[b] NT = not tested.

[c] Not included in any mean as WB control sediment was not tested.

[d] N-1, Mean-1, SD-1 and CV-1 include all data (except Laboratory C) whether control met acceptable limits or not in WB sediment.

[e] N-2, Mean-2, SD-2 and CV-2 include only data for sediment samples from laboratories that met the control performance acceptability criteria in WB sediment.

Table 17.10 Interlaboratory Precision for Survival (%) of *Chironomus tentans* in 10-d Whole-sediment Toxicity Tests (1996/1997)

	Mean Percent Survival (SD) in Sediment Samples and Cd-Spiked Control Sediment													
	Sediment								Cadmium -FS Spikes (mg/kg)					
Laboratory	WB		DC		LS		FS		0.3-Cd		1-Cd		3-Cd	
A	81	(13.6)	79	(6.4)	NT[a]		88	(10.35)	NT		NT		NT	
B	100	(0)	89	(9.1)	93	(8.9)	90	(7.56)	NT		NT		NT	
C	NT		NT		NT		98[b]	(5.00)	98[b]	(5.0)	95[b]	(5.8)	85[b]	(19.1)
E	94	(7.4)	93	(11.7)	84	(13.0)	96	(5.18)	83	(17.1)	85	(5.8)	73	(9.6)
F	99	(3.5)	84	(10.6)	84	(7.4)	88	(8.86)	95	(5.8)	93	(9.6)	98	(5.0)
G	85	(10.7)	76	(20.7)	19	(27.5)	74	(24.46)	NT		NT		NT	
H	96	(7.4)	93	(7.1)	94	(7.4)	100	(0)	NT		NT		NT	
I	90	(7.6)	83	(13.9)	74	(10.6)	86	(14.08)	85	(12.9)	93	(9.6)	83	(15.0)
J	38[c]	(25.5)	25[c]	(20.7)	83[c]	(13.9)	48[c]	(35.76)	23[c]	(22.2)	63[c]	(28.7)	40[c]	(24.5)
K	96	(5.2)	84	(10.6)	NT		98	(4.63)	NT		95	(10.0)	NT	
L	84	(13.0)	70	(13.1)	86	(11.9)	86	(13.02)	NT		NT		NT	
N	83	(12.8)	46	(32.9)	86	(11.9)	91	(17.27)	88	(12.6)	95	(5.8)	70	(8.2)
O	51[c]	(21.0)	61[c]	(18.1)	91[c]	(8.4)	51[c]	(14.58)	85[c]	(5.8)	90[c]	(8.2)	95[c]	(5.8)
P	78	(10.4)	70	(17.7)	41	(24.2)	88	(13.89)	93	(9.6)	93	(9.6)	73	(9.6)
Q	91	(8.4)	93	(8.9)	94	(11.9)	99	(3.54)	98	(5.0)	98	(5.0)	98	(5.0)
R	82	(3.4)	71	(15.4)	56	(13.2)	77	(5.89)	81	(8.0)	83	(11.8)	72	(29.4)
S	75	(14.1)	75	(27.8)	60	(15.1)	71	(18.08)	NT		NT		NT	
X	100	(0)	89	(12.5)	51	(21.7)	98	(7.07)	NT		NT		NT	
N-1[d]	17		17		15		17		9		10		9	
Mean-1	84		75		68		84		81		89		78	
SD-1	16.9		18.0		27.9		15.5		22.6		10.2		18.4	
CV-1	20.0		23.9		41.0		18.5		27.8		11.4		23.6	
N-2[e]	15		15		13		15		7		8		7	
Mean-2	89		80		71		89		89		92		81	
SD-2	8.3		12.5		23.6		9.1		6.5		5.2		12.3	
CV-2	9.4		15.7		33.3		10.2		7.3		5.6		15.2	

[a] NT = not tested.
[b] Not included in any mean as WB control sediment was not tested.
[c] Control survival below acceptable level of 70% in WB sediment.
[d] N-1, Mean-1, SD-1 and CV-1 include all data (except Laboratory C) whether control met acceptable limits or not in WB sediment.
[e] N-2, Mean-2, SD-2 and CV-2 include only data for sediment samples from laboratories that met the control performance acceptability criteria in WB sediment.

Table 17.11 Interlaboratory Precision for Growth (mg/Individual dry weight) of *Chironomus tentans* in 10-d Whole-sediment Toxicity Tests (1996/1997)

	Mean Growth as Dry Weight (SD) in Sediment Samples and Cd-Spiked Control Sediment													
	Sediment								Cadmium -FS Spikes (mg/kg)					
Laboratory	WB		DC		LS		FS		0.3-Cd		1-Cd		3-Cd	
A	0.94	(0.15)	0.38	(0.09)	NT[a]		1.22	(0.27)	NT		NT		NT	
B	1.02	(0.06)	0.24	(0.03)	0.90	(0.34)	1.37	(0.12)	NT		NT		NT	
C	NT		NT		NT		0.86[b]	(0.12)	0.83[b]	(0.14)	0.83[b]	(0.14)	0.20[b]	(0.09)
E	2.47	(0.30)	1.05	(0.21)	2.69	(0.42)	2.29	(0.51)	3.44	(0.29)	2.42	(0.41)	2.90	(0.58)
F	1.69	(0.17)	0.41	(0.13)	1.62	(0.29)	2.43	(0.40)	2.48	(0.26)	2.50	(0.29)	1.02	(0.43)
H	0.92	(0.12)	0.24	(0.05)	0.93	(0.06)	1.29	(0.21)	NT		NT		NT	
I	1.55	(0.27)	0.37	(0.17)	1.80	(0.40)	1.74	(0.49)	2.58	(0.25)	2.05	(0.57)	2.05	(0.50)
J	0.90[c]	(0.83)	0.15[c]	(0.06)	0.91[c]	(0.69)	0.36[c]	(0.23)	1.02[c]	(0.87)	0.42[c]	(0.25)	0.18[c]	(0.05)
K	1.48	(0.12)	0.20	(0.03)	NT		1.68	(0.18)	NT		1.29	(0.05)	NT	
N	0.22[d]	(0.11)	0.06[d]	(0.02)	0.30[d]	(0.06)	0.32[d]	(0.10)	0.35[d]	(0.17)	0.27[d]	(0.04)	0.12[d]	(0.02)
O	0.99[c]	(0.17)	0.07[c]	(0.03)	0.81[c]	(0.07)	1.37[c]	(0.29)	0.67[c]	(0.09)	0.55[c]	(0.06)	0.15[c]	(0.02)
P	1.36	(0.18)	1.01	(0.21)	0.87	(0.31)	0.99	(0.29)	1.63	(0.68)	1.54	(0.18)	1.11	(0.03)
Q	1.01	(0.29)	0.21	(0.09)	1.31	(0.27)	1.08	(0.17)	1.06	(0.15)	1.16	(0.18)	1.16	(0.10)
R	1.31	(0.29)	0.58	(0.28)	1.06	(0.36)	1.51	(0.34)	1.25	(0.38)	1.37	(0.28)	0.70	(0.24)
S	1.73	(0.29)	0.48	(0.21)	2.36	(0.35)	1.26	(0.80)	NT			NT	NT	
X	0.97	(0.10)	0.68	(0.14)	0.95	(0.36)	1.09	(0.22)	NT			NT	NT	
N-1[e]	15		15		13		15		9		10		9	
Mean-1	1.24		0.41		1.27		1.33		1.61		1.36		1.04	
SD-1	0.51		.31		0.67		0.58		1.02		0.80		0.93	
CV-1	41.6		75.3		53.1		43.3		63.3		58.7		69.6	
N-2[f]	12		12		10		12		6		7		6	
Mean-2	1.39		0.49		1.45		1.50		2.1		1.76		1.49	
SD-2	0.45		0.29		0.65		0.47		0.92		0.56		0.83	
CV-2	33.2		60.2		45.1		31.1		44.2		31.5		55.3	

[a] NT = not tested.

[b] Not included in any mean as WB control sediment was not tested.

[c] Control survival below acceptable level of 70% in WB sediment.

[d] Control weight below acceptable level of 0.60 mg/organism in WB sediment.

[e] N-1, Mean-1, SD-1 and CV-1 include all data (except Laboratory C) whether control met acceptable limits or not in WB control sediment.

[f] N-2, Mean-2, SD-2 and CV-2 include only data for sediment samples from laboratories that met the control performance acceptability criteria in WB sediment.

Table 17.12 Interlaboratory Precision for Growth (mg/Individual as ash-free dry weight) of *Chironomus tentans* in 10-d Whole-sediment Toxicity Tests (1996/1997)

	Mean Growth as Ash-free Dry Weight (SD) in Sediment Samples and Cd-spiked Control Sediment													
	Sediment								Cadmium -FS Spikes (mg/kg)					
Laboratory	WB		DC		LS		FS		0.3-Cd		1-Cd		3-Cd	
B	0.79	(0.03)	0.18	(0.03)	0.69	(0.07)	1.04	(0.09)	NT[a]		NT		NT	
C	NT		NT		NT		0.20[b]	(0.05)	0.19[b]	(0.03)	0.23[b]	(0.12)	0.03[b]	(0.03)
E	0.25[c]	(0.09)	0.10[c]	(0.03)	0.2[c]	(0.07)	0.24[c]	(0.06)	0.48[c]	(0.12)	0.27[c]	(0.08)	0.38[c]	(0.18)
F	0.50	(0.11)	0.13	(0.12)	0.73	(0.16)	1.14	(0.39)	0.94	(0.10)	1.00	(0.31)	0.45	(0.24)
I	1.35	(0.26)	0.32	(0.13)	1.16	(0.27)	1.99	(1.50)	2.01	(0.19)	1.56	(0.35)	1.55	(0.41)
K	1.06	(0.09)	0.17	(0.02)	NT		1.12	(0.09)	NT		0.91	(0.03)	NT	
L	1.07	(0.28)	0.34	(0.09)	1.13	(0.23)	1.11	(0.18)	NT		NT		NT	
O	0.30[c,d]	(0.05)	0.01[c,d]	(0.01)	0.26[c,d]	(0.06)	0.60[c,d]	(0.15)	0.22[c,d]	(0.03)	0.16[c,d]	(0.03)	0.03[c,d]	(0.01)
P	0.36[d]	(0.33)	0.29[d]	(0.03)	0.18[d]	(0.10)	0.15[d]	(0.05)	0.46[d]	(0.41)	0.29[d]	(0.07)	0.21[d]	(0.05)
Q	0.76	(0.24)	0.15	(0.08)	0.78	(0.16)	0.79	(0.12)	0.74	(0.12)	0.78	(0.22)	0.78	(0.04)
R	0.88	(0.27)	0.40	(0.16)	0.64	(0.17)	0.94	(0.20)	0.74	(0.21)	0.86	(0.22)	0.46	(0.17)
X	0.15[d]	(0.04)	0.20[d]	(0.09)	0.49[d]	(0.21)	0.30[d]	(0.18)	NT		NT		NT	
N-1[e]	11		11		10		11		7		8		7	
Mean-1	0.677		0.208		0.630		0.856		0.799		0.729		0.551	
S D-1	0.39		0.12		0.35		0.53		0.58		0.47		0.50	
CV-1	58.1		56.1		54.9		61.8		73.1		64.6		90.2	
N-2[f]	7		7		6		7		4		5		4	
Mean-2	0.916		0.241		0.855		1.161		1.108		1.022		0.810	
SD-2	0.27		0.11		0.23		0.39		0.61		0.31		0.52	
CV-2	29.8		45.0		26.8		33.2		55.0		30.4		63.8	

[a] NT = not tested.

[b] Not included in any mean as WB control sediment was not tested.

[c] Control weight below acceptable weight criteria of 0.48 mg/organism in WB sediment.

[d] Control survival below acceptable level of 70% in WB sediment.

[e] N-1, Mean-1, SD-1 and CV-1 include all data (except Laboratory C) whether control met acceptable limits or not in WB sediment.

[f] N-2, Mean-2, SD-2 and CV-2 include only data for sediment samples from laboratories that met the control performance acceptability criteria in WB sediment.

of MDDs for AFDW across laboratories was 0.03 to 0.22 mg for the DC sediment and 0.04 to 0.25 mg for LS sediment. No toxicity relative to weight was observed in the cadmium tests. The mean dry weight of midge in the 1 mg/kg Cd treatment was 1.76 mg/organism (CV=32%). The mean MDD and range for the 1 mg/kg Cd sample was 0.28 mg/organism (0.09 to 0.57). The AFDW for the 1 mg/kg sample was 1.022 mg/organism (CV=30%) with MDDs of 0.19 mg (0.04 to 0.36).

17.5.4 These round-robin tests conducted in 1993 (Section 17.5.2) and in 1996/1997 (Section 17.5.3) exhibited similar or better precision compared to many chemical analyses and effluent toxicity test methods (USEPA, 1991a; USEPA, 1991c). The success rate for test initiation and completion of the USEPA's round-robin evaluations is a good indication that a well equipped and trained staff will be able to successfully conduct these tests. This is an important consideration for any test performed routinely in any regulatory program.

17.6 Precision of Sediment Toxicity Test Methods: Evaluation of Long-term Sediment Tests

17.6.1 Interlaboratory precision evaluations of the long-term *H. azteca* and *C. tentans* tests, using the methods described in Sections 14 and 15, were conducted by federal government, contract, and academic laboratories that had demonstrated experience in sediment toxicity testing, although only two of the laboratories had prior experience with the long-term test methods described in this manual. This round-robin study was conducted in two phases: a Preliminary Round-robin (PRR) and a Definitive Round-robin (DRR). The objective of the PRR was to provide participating laboratories with an opportunity to become acquainted with the techniques necessary to conduct the two tests and to solicit commentary and recommendations regarding potential improvements for the definitive evaluation. Criteria for selection of participants in both phases were that the laboratories: (1) had existing cultures of the test organisms, (2) had experience conducting 10-d tests with the organisms, and (3) would participate voluntarily. Methods for conducting toxicity tests were similar among laboratories, and each laboratory was supplied with detailed operating procedures outlining these methods. Methods for culturing were not specified and were not identical across laboratories (as long as each laboratory started with the appropriate age test organisms). The PRR (phase 1) included the WB control sediment (West Bearskin, MN; WB) and the formulated sediment (FS) in which alpha-cellulose represented the primary carbon source (Kemble et al., 1999; Table 17.13). The DRR (phase 2) also included a copper-contaminated sediment from Cole Creek, Keweenaw, MI (CC), and a PAH-contaminated sediment from the Little Scioto River, OH (LS). In addition to the WB control sediment and the FS sediment described above, an additional sediment, in which peat (PE) represented the primary carbon source, was also tested (Table 17.13).

17.6.2 Twelve laboratories participated in the PRR with *H. azteca*. In these tests, 100% of laboratories passed the acceptability criterion for survival (≥80%) in the WB control sediment at 28 d (Table 17.14) with survival ranges of 83 to 98% at 28 d, 71 to 93% at 35 d and 63 to 92% at 42 d. In the formulated sediment (FS), 80% of the laboratories met the survival criterion at 28 d (range: 47 to 98%). Survival ranges in FS sediment at 35 d were 48 to 98% and at 42 d the survival ranges were 48 to 98%. For growth measured as length in the WB sediment, 92% of the laboratories reported the mean length of the organisms to be ≥3.2 mm at 28 d (range: 3.07 to 5.64 mm). For the FS sediment, 100% of the laboratories reported length ≥3.2 mm with lengths ranging from 3.54 to 5.44 mm. For growth measured as dry weight, >66% of the laboratories met the minimum weight criterion (≥0.15 mg/organism) in WB (range: 0.10 to 1.16 mg/individual). In the FS samples, 100% of the laboratories met this growth criterion, with weight ranges from 0.15 to 0.90 mg/individual. The criterion for reproductive output for *H. azteca* (≥2 young/female) was met by 78% of laboratories in the WB (range: 0 to 27 young/female). In the FS samples, 89% of the laboratories met the reproductive requirement with ranges of 0.62 to 22 young/female.

17.6.3 Ten laboratories participated in the PRR with *C. tentans*. In these tests, 90% of laboratories passed the acceptability criterion for survival at 20 d (≥70%) in WB (range: 67 to 96%; Table 17.14), and in the FS sediment, 60% of the laboratories met the acceptability criterion (range: 42 to 83%). For growth measured as dry weight, 100% of laboratories passed the criterion (≥0.6 mg/individual) in WB (range: 1.45 to 3.78 mg/individual). For the FS samples, 86% of the laboratories passed the criterion (range: 0.50 to 3.40 mg/individual). For growth as AFDW, 100% of the laboratories passed the criterion of ≥0.48 mg in the WB (range: 0.86 to 3.22 mg/individual) (Table 17.14). In the FS sediment, 88% of the laboratories met the growth criterion (as dry weight) with ranges of weights from 0.42 to 2.72 mg/individual. The criterion for emergence (≥50%) was met by 70% of the laboratories in WB sediment. In the FS, 50% of the laboratories met the emergence criterion. The criterion for reproductive output in *C. tentans* (≥800 eggs/female) was exceeded by 90% of laboratories in WB control sediment (range: 504 to 1240 eggs/female). In FS, 86% of laboratories met this criterion in the FS (range: 0 to 1244 eggs/female). The suggested criterion for percent hatch (≥80%) was met by 88% of laboratories in WB (range: 0 to 98%), and in FS, 67% of laboratories (range: 0 to 98.7%).

17.6.4 In both the *H. azteca* and *C. tentans* tests, the results of the PRR demonstrated that the majority of laboratories met the acceptability criteria for those endpoints for which criteria had been established (e.g., survival and growth). The highest proportion of failures in the midge test occurred with post-pupation endpoints (emergence, percent hatch) and may reflect the fact that the criteria developed for these endpoints are based on evaluations conducted at a single laboratory (Sibley et al., 1996; Sibley et al., 1997b; Benoit et al., 1997). In the PRR, some laboratories experienced unacceptably low oxygen

Table 17.13 Physical Characteristics of the Sediments Used in the Preliminary and Definitive Round-robin Evaluations of Long-term Methods for Sediment Toxicity Testing (Section 17.6).

Sediment	Total Organic Carbon (%)	Water Content	Particle Size (%) Sand	Clay	Silt	Sediment Type
FS[a] (a high sand/low TOC)	2.2	31	74	16	11	Sandy Loam
WB	3.3	31	74	16	10	Sandy Loam
PE	10	ND[a]	ND	ND	ND	Clay

[a] ND = not determined

Table 17.14 Percentage of Laboratories Meeting Performance Levels for the Following Endpoints in the WB Control Sediment Evaluated in the Long-term Round-robin Tests.

Performance Level	Preliminary Round	Definitive Round
	Hyalella azteca	
28-d survival ≥ 80%	100	88
28-d growth ≥ 3.2 mm length	92	71
28-d growth ≥0.15 mg dry weight	66	88
28- to 42-d reproduction (≥ 2 young/female)	78	71
	Chironomus tentans	
20-d survival ≥70%	90	63
20-d growth ≥0.6 mg (dry weight)	100	63
20-d growth ≥0.48 mg (ash-free dry weight)	100	67
Emergence ≥50%	70	50
Number of eggs/egg case ≥ 800	90	63
Percentage hatch ≥80%	88	57

Table 17.15 **Interlaboratory Comparison of Day 28 Percent Survival (Mean ± SD) of *H. azteca* in a Long-term Sediment Exposure Using Five Sediments (WB = West Bearskin, CC = Cole Creek, LS = Little Scioto River, FS = Formulated Sediment (using alpha-cellulose as organic carbon source), and PE = Formulated Sediment (using peat moss as organic carbon source)).**

	Sediment									
Laboratory	WB		CC		LS		FS		PE	
E	100	(0)	97	(4.9)	94	(6.7)	94	(7.9)	NT [a]	
F	62[b]	(33.0)	84[b]	(21.1)	90[b]	(20.9)	38[b]	(35.2)	93[b]	(23.0)
H	93	(9.6)	85	(5.8)	98	(5.0)	NT		68	(37.8)
K	95	(10.0)	98	(4.5)	96	(6.7)	NT			NT
L	83	(12.2)	88	(8.7)	84	(12.4)	78	(14.2)	54	(40.6)
N	89	(16.8)	92	(8.4)	91	(6.7)	NT		NT	
Q	98	(5.0)	93	(9.6)	80	(27.1)	90	(14.4)	88	(9.6)
U	100	(0)	100	(0)	98	(5.0)	NT		NT	
X	NT		NT		NT		83[c]	(10.7)	93[c]	(7.5)
N-1[d]	8		8		8		4		4	
Mean-1	90		92		91		75		75	
SD-1	12.8		6.0		6.4		25.6		17.8	
CV (%)-1	14.3		6.5		7.0		34.6		23.6	
N-2[e]	7		7		7		3		3	
Mean-2	94		93		91		87		70	
SD-2	6.4		5.5		6.9		8.7		16.8	
CV (%)- 2	6.8		5.9		7.5		9.9		24.1	

[a] NT = not tested

[b] Control survival below acceptable level of 80% in WB sediment.

[c] Not included in any mean as WB control sediment was not tested.

[d] N-1, Mean-1, SD-1 and CV-1 include all data (except Laboratory X) whether control met acceptable limits or not in WB sediment.

[e] N-2, Mean-2, SD-2 and CV-2 include only data for sediment samples from laboratories that met the 28-d control performance acceptability criteria in WB sediment.

levels during evaluation of the *C. tentans* test which was attributed to high feeding rates. To address this issue, the feeding rate for the DRR of the *C. tentans* test was reduced from 1.5 to 1.0 mL/d of Tetrafin.

17.6.5 In total, eight laboratories participated in the DRR with *H. azteca*; however all laboratories did not test all sediments. Mean survival for those laboratories that met the control survival test acceptability criteria at 28 d in the WB control sediment was 94% (CV=6.8%). In FS, the mean survival was 87% (CV=9.9%), and in the PE it was 70% (CV=24%; Table 17.15). Mean survival at 35 d with laboratories that met the ≥80% control survival criterion at 28 d was as follows: WB had 92% survival (CV=7.2%), FS had 88% survival (CV=15.1%) and PE had survival of 63% (CV=34.0%; Table 17.16). Mean survival at 42 d with laboratories that met the ≥80% 28-d control survival criterion was as follows: WB had 92% survival (CV=7.4%), FS had 84% survival (CV=14.1%) and PE had 60% survival (CV=38.2% with 3 laboratories; Table 17.16). At 28 d, 88% of the laboratories met the control survival criteria in the WB control sediment (Table 17.14). When acceptable 28-d control survival was reported in WB sediment, 71% of the laboratories met the length criterion (≥3.2 mm) for *H. azteca* (Table 17.14). For those laboratories that met the 28-d survival criterion and the growth criterion, the mean growth (measured as length) of *H. azteca* at 28 d was 4.17 mm (CV=12.4%) in WB, 3.51 mm (CV=22.6%) in the FS and 3.24 mm (CV=36.6%) in the PE (Table 17.18). For growth measured as dry weight for the WB control sediment, 88% of the laboratories met the weight criterion of ≥0.15 mg/individual when acceptable 28-d control survival was reported (Table 17.19) The mean growth of *H. azteca* (mg/individual dry weight) in each sample where 28-d control survival and growth was met was: 0.25 mg (CV=27.8%) in WB, 0.30 mg (CV=68.6%) in FS, and 0.18 mg (CV=34.0%; Table 17.19) in PE. For the WB control sediment, 71% of the laboratories met the reproduction criteria (≥2 young/female) when acceptable 28-d control survival was reported (Table 17.14). The mean reproduction from 28 to 42 d for laboratories that met both the reproduction criteria and 28-d survival criteria was 3.13 young/female (CV=48.9%) for WB. For the FS, only one laboratory that had acceptable survival in WB control sediment at 28 d also had acceptable reproduction at 42 d, with a mean of 2.3 young/female. For the PE sediment, the only laboratory that had acceptable survival did not have acceptable young production, as only 0.08 young/female were obtained (Table 17.20).

Table 17.16 **Interlaboratory Comparison of Day 35 Percent Survival (Mean ± SD) of *H. azteca* in a Long-term Sediment Exposure Using Five Sediments (WB = West Bearskin, CC = Cole Creek, LS = Little Scioto River, FS = Formulated Sediment (using alpha-cellulose as organic carbon source), and PE = Formulated Sediment (using peat moss as organic carbon source)).**

	Sediment									
Laboratory	WB		CC		LS		FS		PE	
E	98	(7.1)	96	(5.2)	96	(5.2)	98	(4.6)	NT[a]	
F	70[b]	(31.6)	73[b]	(31.5)	86[b]	(27.2)	33[b]	(38.5)	86[b]	(35.0)
H	95	(7.6)	96	(7.4)	95	(5.4)	NT		46	(26.7)
K	91	(18.1)	96	(5.2)	90	(10.7)	NT		NT	
L	78	(10.4)	83	(8.9)	84	(13.0)	73	(17.5)	57	(37.4)
N	93	(7.1)	88	(11.7)	83	(7.6)	NT		NT	
Q	94	(9.2)	86	(27.7)	88	(13.9)	93	(8.9)	88	(12.8)
U	95	(7.6)	98	(4.6)	86	(10.6)	NT		NT	
X	NT		NT		NT		74[c]	(16.9)	95[c]	(5.4)
N-1[d]	8		8		8		4		4	
Mean-1	89		89		88		74		69	
SD-1	9.8		8.9		4.9		21.0		30.0	
CV (%)1	11.0		10.0		5.6		30.4		40.1	
N-2[e]	7		7		7		3		3	
Mean-2	92		92		89		88		63	
SD-2	6.6		6.2		5.2		13.2		21.5	
CV (%) -2	7.2		6.7		5.9		15.1		34.0	

[a] NT = not tested

[b] Control survival below acceptable level of 80% in WB sediment at 28 d.

[c] Not included in any mean as WB control sediment was not tested.

[d] N-1, Mean-1, SD1 and CV (%)-1 include all data (except Laboratory X) whether control met acceptable limits or not in WB sediment.

[e] N-2, Mean-2, SD-2 and CV-2 include only data for sediment samples from laboratories that met the 28-d control survival performance acceptability criteria in WB sediment.

17.6.6 Overall, nine laboratories participated in the DRR with *C. tentans* but not all laboratories tested all sediments. Mean survival (with CV in parentheses) for those laboratories that met the control criterion of ≥80% survival at 20 d was 85% (CV=5%) for WB sediment. In addition, mean survival at 28 d, in the FS was 86% (CV=14.4%) and, in the PE sediment was 75% (CV=13.9%) (Table 17.21). In total, 63% of the laboratories met the acceptability criterion for survival (≥70%) for the WB control sediment in the *C. tentans* test (Table 17.14). For laboratories reporting dry weights, the mean growth of *C. tentans* at 20 d (criterion of ≥0.60 mg/individual dry weight and ≥70% survival) was 1.45 mg (CV=58.6%) for WB sediment. In addition, mean growth (as dry weight) was 1.63 mg/individual (CV=20.9%) for the FS and 1.43 mg/individual (CV=47.9%) for the PE sediment (Table 17.22). For laboratories reporting weights as AFDW, the mean growth of *C. tentans* at 20 d (criterion of ≥0.48 mg/individual AFDW and ≥70% survival) was 0.81 mg (CV=53.3%) for WB, 1.05 mg/individual (CV=18.1%) for FS, and 0.64 mg/individual (CV=12.7%) for PE (Table 17.23). For growth as dry weight in the WB control sediment, 63% of the laboratories met the acceptability criterion for survival and growth (as dry weight) in the *C. tentans* test, while for AFDW, 67% of the laboratories met the test acceptability criterion of ≥0.48 mg/AFDW per individual (Table 17.14). Mean percent emergence for those laboratories that met the emergence criterion of ≥50% reported emergence in WB control sediment as 69.8% (CV=29.5%). In addition, mean emergence was 50.5% in FS (CV=68.6%) and 55.8% in PE (CV=30.3%) sediment (Table 17.24). In total 50% of the laboratories met the acceptability criterion for both 20-d survival and emergence in the WB control sediment (Table 17.14). The success rate for the number of eggs /case and the control survival criterion was 63% in WB. Mean number of eggs/female was 1118 eggs/case (CV=15.0%) in WB. The FS and PE sediments had 1024 eggs/case (CV=30.4%) and 867 eggs/case (CV=29.3%), respectively (Table 17.25). The mean percent hatch for laboratories with acceptable control survival and acceptable number of eggs/case was 90% (CV=10.8%) for WB control sediment (Table17.26), and 57% of the laboratories that tested these

Table 17.17 **Interlaboratory Comparison of Day 42 Percent Survival (Mean ± SD) of *H. azteca* in a Long-term Sediment Exposure Using Five Sediments (WB = West Bearskin, CC = Cole Creek, LS = Little Scioto River, FS = Formulated Sediment (using alpha-cellulose as organic carbon source), and PE = Formulated Sediment (using peat moss as organic carbon source)).**

	Sediment									
Laboratory	WB		CC		LS		FS		PE	
E	95	(7.6)	93	(7.1)	95	(5.4)	93	(8.9)	NT[a]	
F	61[b]	(31.8)	68[b]	(33.7)	85[b]	(26.7)	30[b]	(37.6)	83[b]	(33.7)
H	90	(9.3)	90	(9.3)	93	(8.9)	NT		40	(26.2)
K	91	(18.1)	96	(5.2)	88	(12.8)	NT		NT	
L	75	(10.7)	83	(8.9)	84	(13.0)	70	(16.0)	55	(36.7)
N	89	(8.4)	81	(17.3)	79	(10.7)	NT		NT	
Q	93	(11.7)	81	(30.9)	88	(13.9)	89	(13.6)	85	(16.0)
U	93	(8.9)	95	(5.4)	86	(10.6)	NT		NT	
X	NT		NT		NT		43[c]	(23.2)	84[c]	(9.2)
N-1[d]	8		8		8		4		4	
Mean-1	86		86		87		70		65	
SD-1	11.7		9.6		5.1		28.6		21.8	
CV (%)-1	13.6		11.2		5.8		40.7		33.3	
N-2[e]	7		7		7		3		3	
Mean-2	92		88		87		84		60	
SD-2	6.6		6.6		5.4		12.1		22.9	
CV (%)2	7.4		7.5		6.2		14.1		38.2	

[a] NT = not tested

[b] Control survival below acceptable level of 80% in WB sediment at 28 d.

[c] Not included in any mean as WB control sediment was not tested.

[d] N-1, Mean-1, SD1 and CV (%)-1 include all data (except Laboratory X) whether control met acceptable limits or not in WB sediment.

[e] N-2, Mean-2, SD-2 and CV (%)-2 include only data for sediment samples from laboratories that met the 28-d control survival performance acceptability criteria in WB sediment.

sediments met the test acceptability criteria for hatchability.

17.6.7 In total, the proportion of laboratories that met the various endpoint criteria in WB control sediment in the DRR was higher for *H. azteca* than it was for *C. tentans*. The most likely reason for the lower success with *C. tentans* in the DRR was the reduction in feeding rate (from 1.5 to 1.0 ml of Tetrafin/beaker/d) relative to the PRR. In the PRR with *C. tentans*, the proportion of laboratories meeting the various endpoint criteria was generally higher (see Table 17.14), particularly for post-pupation endpoints (emergence, reproduction, and percent hatch). Therefore, this manual recommends that the higher feeding rate of 1.5 ml/beaker/d be used in long-term tests with *C. tentans* (Section 15).

17.6.8 In the DRR, mean survival (CV in parentheses) of *H. azteca* in the LS sediment (contaminated with PAHs; using only values where the 28-d control survival criterion was met) was 91% (CV=7.5%) at 28 d, was 89% (CV=5.9%) at 35 d and 87% (CV=6.2%) at 42 d (Tables 17.15 to 17.17). Mean survival of *C. tentans* at 20 d in the LS sediment was 40% (CV=82.6%; Table 17.21). The growth of *H. azteca* in LS sediment resulted in a mean length of 4.37 mm (CV=10.1%; Table 17.18) and a mean dry weight of 0.31 mg/individual (CV=38.2%; Table 17.19). Mean growth of *C. tentans* in LS was 1.72 mg/individual (CV=66.2%) as dry weight (Table 17.22) and 2.31 mg/individual (CV=59.1%) as AFDW (Table 17.23). For both species, all growth endpoints were highest for LS relative to the other sediments evaluated, except for *H. azteca* dry weight which had a comparable mean as the other four sediments. The mean proportion of *C. tentans* larvae emerging from LS was 35.7% (CV=71.2%; Table 17.24). This value was roughly half of the emergence from the control sediments. Mean reproductive output of *H. azteca* in LS sediment, for those laboratories with acceptable control survival, was 3.08 young/female (CV of 41.0%; Table 17.20). The mean reproductive output of *C. tentans* in the LS sediment for laboratories that met the control survival criteria was 980 eggs/female (CV=20.1%; Table 17.25), which was similar to the WB, FS, and PE sediments. Mean percent hatch of *C. tentans* eggs was 94% (CV=6.5%) for the laboratories that met at least 70% control survival (Table 17.26).

Table 17.18. **Interlaboratory Comparison of Day 28 Length (Mean mm/Individual ± SD) of *H. azteca* in a Long-term Sediment Exposure Using Five Sediments (WB = West Bearskin, CC = Cole Creek, LS = Little Scioto River, FS = Formulated Sediment (using alpha-cellulose as organic carbon source), and PE = Formulated Sediment (using peat moss as organic carbon source)).**

	Sediment									
Laboratory	WB		CC		LS		FS		PE	
E	4.15	(0.23)	4.00	(0.11)	4.29	(0.16)	2.96	(0.03)	NT[a]	
F	3.02 [b,c]	(0.28)	4.66 [b,c]	(0.17)	5.23 [b,c]	(0.41)	3.70 [b,c]	(0.30)	5.03[b,c]	(0.06)
H	3.77	(0.32)	2.72	(0.14)	3.77	(0.17)	NT		2.40	(0.41)
K	4.18	(0.12)	4.39	(0.29)	4.95	(0.22)	NT		NT	
L	5.02	(0.11)	4.97	(0.27)	4.62	(0.40)	4.07	(0.39)	4.08	(0.64)
N	NR		NR		NR		NR		NR	
Q	3.11[c]	(0.10)	3.17[c]	(0.18)	4.29[c]	(0.45)	4.51[c]	(0.46)	3.27[c]	(0.03)
U	3.74	(0.08)	3.99	(0.17)	4.21	(0.13)	4.21	NA	NT	
X	NT		NT			NT	3.25[d]	(0.20)	3.35[d]	(0.21)
N-1[e]	7		7		7		5		4	
Mean-1	3.86		3.99		4.40		3.81		3.69	
SD-1	0.68		0.76		0.50		0.66		0.99	
CV (%)-1	17.8		20.1		10.9		17.3		30.4	
N-2[f]	5		5		5		2		2	
Mean-2	4.17		4.01		4.37		3.51		3.24	
SD-2	0.52		0.83		0.44		0.79		1.19	
CV (%)-2	12.4		20.6		10.1		22.6		36.6	

[a] NT = not tested; NR = not reported; NA = not applicable.

[b] Control survival below acceptable level of 80% in WB sediment at 28 d.

[c] Length below acceptable level of 3.2 mm in length in WB control sediment.

[d] Not included in any mean as WB control sediment was not tested.

[e] N-1, Mean-1, SD1 and CV (%)-1 include all data (except Laboratory X) whether control met acceptable limits or not in WB sediment.

[f] N-2, Mean-2, SD-2 and CV-2 include only data for sediment samples from laboratories that met the 28-d control survival performance acceptability criteria in WB sediment.

17.6.9 Across all laboratories that met the 28-d survival criterion of ≥80% for *H. azteca*, the mean survival in the contaminated CC sediment sample was 93% (CV=5.9%) at 28 d, 92% (CV=7.2%) at 35 d, and 88% at 42 d (CV=7.5%; Tables 17.15 to 17.17). Mean survival of *C. tentans* at 20 d for laboratories that met the 20-d control survival criteria was 75% (CV=30.9%; Table 17.21). In CC sediment, the mean growth of *H. azteca* was 4.01 mm (CV=20.6%) as length (Table 17.18) and 0.24 mg/individual (CV=75.2%) as dry weight (Table 17.19). Mean growth of *C. tentans* in CC sediment was 0.68 mg/individual (CV=66.0%) as dry weight (Table 17.22) and 0.37 mg/individual (CV=49.6%) as AFDW (Table 17.23). The growth was reduced about 50% in the CC sediment in comparison to the WB, FS, and PE sediments for *C. tentans* only. The mean proportion of *C. tentans* larvae to emerge from CC sediment was 38% (CV=60.5%; Table 17.24). Similar to the LS sediment sample, this emergence was reduced to about half of that observed in the control sediments. Mean reproductive output of *H. azteca* in CC sediment, for those laboratories with acceptable 28-d control survival, was 1.64 young/female (CV=103.3%) in contrast to the mean for WB of 3.13 young/female (CV=48.9%; Table 17.20). The mean reproductive output of *C. tentans* eggs in the CC sediment for laboratories that met the 20-d control survival criteria was 621 eggs/female (CV=52.4%) (Table 17.25) which was the lowest egg production for all sediments, which averaged between 404-1194 eggs/female. The mean percent hatch of *C. tentans* eggs was 69% (CV=49.5%) for the laboratories that met at least 70% control survival (Table 17.26); all other sediments had percent hatches for survival averaging 90 to 94%.

17.6.10 For the chronic *H. azteca* test, the mean MDD for survival relative to the WB control sediment for the CC sediment across all laboratories was only 7.7% (2.4 to 19.5%) at 28 d and 12.8% (6.4 to 28l.7%) at day 42. The MDDs for survival of amphipods were also small in the LS sediment: 10.8% (3.3 to 26%) at 28 d and 11.5% (5.7 to 26%) at 42 d. The mean MDDs relative to WB control sediment were also low for the 28-d amphipod weights as the mean MDD for the CC sediment relative to WB control sediment was 0.06 mg (0.04 to 0.14 mg) and the mean MDD

Table 17.19. Interlaboratory Comparison of Day 28 Dry Weight (Mean mg/Individual ± SD) of *H. azteca* in a Long-term Sediment Exposure Using Five Sediments (WB = West Bearskin, CC = Cole Creek, LS = Little Scioto River, FS = Formulated Sediment (using alpha-cellulose as organic carbon source), and PE = Formulated Sediment (using peat moss as organic carbon source)).

	Sediment									
Laboratory	WB		CC		LS		FS		PE	
E	0.29	(0.04)	0.23	(0.02)	0.34	(0.07)	0.12	(0.02)	NT[a]	
F	0.01[b,c]	(0.01)	0.49[b,c]	(0.04)	0.78[b,c]	(0.18)	0.11[b,c]	(0.15)	0.73[b,c]	(0.10)
H	0.25	(0.06)	0.10	(0)	0.20	(0)	NT		0.15	(0.06)
K	0.31	(0.04)	0.56	(0.05)	0.58	(0.09)	NT		NT	
L	0.36	(0.04)	0.41	(0.07)	0.32	(0.12)	0.40	(0.10)	0.24	(0.05)
N	0.23	(0.10)	0.09	(0.03)	0.25	(0.09)	NT		NT	
Q	0.16	(0.04)	0.09	(0.01)	0.31	(0.09)	0.39	(0.06)	0.13	(0.01)
U	0.19	(0.02)	0.21	(0.03)	0.27	(0.04)	NT		NT	
X	NT		NT		NT		0.22[d]	(0.17)	0.42[d]	(0.37)
N-1[e]	8		8		8		4		4	
Mean-1	0.22		0.27		0.38		0.23		0.31	
SD-1	0.11		0.19		0.20		0.16		0.28	
CV (%)-1	48.8		69.6		52.1		71.2		90.0	
N-2[f]	7		7		7		3		3	
Mean-2	0.25		0.24		0.31		0.30		0.18	
SD-2	0.07		0.18		0.12		0.21		0.06	
CV (%)-2	27.8		75.2		38.2		68.6		34.0	

[a] NT = not tested.

[b] Control survival below acceptable level of 80% in WB sediment at 28 d.

[c] Weight below test acceptable criteria of 0.15 mg/organism in WB control sediment.

[d] Not included in any mean as WB control sediment was not tested.

[e] N-1, Mean-1, SD1 and CV (%)-1 include all data (except Laboratory X) whether control met acceptable limits or not in WB sediment.

[f] N-2, Mean-2, SD-2 and CV-2 include only data for sediment samples from laboratories that met the 28-d control survival performance acceptability criteria in WB sediment.

for length was 0.26 mm (0.18 to 0.33 mm). The mean MDD for LS sediment for amphipod growth as weight was 0.10 mg (0.05 to 0.16 mg) and length of 0.33 mm (0.14 to 0.44 mm). The mean MDD for the mean number of young per female was 1.92 (0.09 to 2.4) in CC sediment and 2.06 (0.57 to 3.1) in LS sediment relative to WB control sediment.

17.6.11 The summary of the MDDs relative to the WB control sediment for CC and LS samples and the chronic *C. tentans* test is discussed by endpoint. For percent survival at 20 d, the mean MDDs relative to WB control sediment for CC and LS sediments were 14.4% (range of 5.9 to 19.1%) and 15.6% (5.8 to 25.3%), respectively. For 20 d dry weights, the mean MDDs were 24.9% (CC) and 64.2% (LS) with ranges of 15.6 to 30.4% and 25.1 to 126.9%, respectively. The mean MDD and range for the AFDW relative to the WB control sediment was 29.9% (22.9 to 44.6%) for the CC sediment and 68.7% (22.9 to 125.0%) for LS sediment. For emergence the mean MDD for the CC sediment was 19.4% (10.5 to 25.0%) and the mean LS MDD was 17.9 (8.2 to 23.0%). The number of eggs produced had a mean MDD relative to the WB control sediment of 19.4% (11.0 to 29.3%) for the CC sediment and 24.4% (11.9 to 37.4%) for LS sediment, while hatch had a mean MDD of 42.2% (7.4 to 77.3%) for the CC sediment and 30.5% for LS sediment (9.3 to 53.7%).

17.6.12 These chronic round-robin tests exhibited similar or better precision compared to many chemical analyses and effluent toxicity test methods (USEPA, 1991a; USEPA, 1991c). The success rate for test initiation and completion of the USEPA's round-robin evaluations is a good indication that a well equipped and trained staff will be able to successfully conduct these tests. These are very important considerations for any test performed routinely in any regulatory program.

Table 17.20 **Interlaboratory Comparison of Reproduction (Mean Number of Young/Female ± SD) of *H. azteca* in a Long-term Sediment Exposure Using Five Sediments (WB = West Bearskin, CC = Cole Creek, LS = Little Scioto River, FS = Formulated Sediment (using alpha-cellulose as organic carbon source), and PE = Formulated Sediment (using peat moss as organic carbon source)).**

	Sediment									
Laboratory	WB		CC		LS		FS		PE	
E	5.7	(3.1)	4.2	(2.2)	4.2	(1.6)	2.3	(2.9)	NT[a]	
F	4.0[b]	(4.7)	7.5[b]	(7.6)	19.4[b]	(4.4)	5.4[b]	(2.1)	16.5[b]	(9.4)
H	2.3	(2.6)	0.3	(0.2)	1.2	(1.3)	NT		0.08	(1.8)
K	3.3	(1.9)	1.2	(1.4)	4.1	(4.5)	NT		NT	
L	NA[a]		NA		NA		NA		NA	
N	2.0	(1.5)	0.2	(0.7)	2.2	(1.3)	NT		NT	
Q	0.09[c]	(0.1)	0.04[c]	(0.04)	0.6[c]	(0.9)	0.2[c]	(0.2)	0.3[c]	(0.4)
U	2.4	(1.5)	2.4	(1.7)	3.5		NT		NT	
X	NT		NT		NT		0.12[d]	(0.73)	0.5[d]	(0.7)
N-1[e]	7		7		7		3		3	
Mean-1	2.8		2.2		5.0		2.6		5.9	
SD-1	1.8		2.7		6.5		2.7		9.2	
CV (%)-1	62.6		121.6		128.2		100.5		157.3	
N-2[f]	5		5		5		1		1	
Mean-2	3.13		1.64		3.08		2.3		0.08	
SD-2	1.53		1.69		1.27		--		--	
CV (%)-2	48.9		103.3		41.0		--		--	

[a] NT = not tested; NA = not applicable: young count not reported per female.

[b] Survival below test acceptable criteria in WB control sediment at 28 d.

[c] Reproduction below test acceptable criteria in WB control sediment of 2 young/female.

[d] Not included in any mean as WB control sediment was not tested.

[e] N-1, Mean-1, SD1 and CV (%)-1 include all data (except Laboratory X) whether control met acceptable limits or not in WB sediment.

[f] N-2, Mean-2, SD-2 and CV-2 include only data for sediment samples from laboratories that met the 28-d control survival performance acceptability criteria in WB sediment.

Table 17.21 **Interlaboratory Comparison of Day 20 Percent Survival (Mean ± SD) of *C. tentans* in a Long-term Sediment Exposure Using Five Sediments (WB = West Bearskin, CC = Cole Creek, LS = Little Scioto River, FS = Formulated Sediment (using alpha-cellulose as organic carbon source), and PE = Formulated Sediment (using peat moss as organic carbon source)).**

	Sediment									
Laboratory	WB		CC		LS		FS		PE	
E	94	(8)	98	(4)	19	(13)	94	(8)	NT[a]	
F	79	(16)	40	(4)	17	(7)	81	(8)	65	(10)
H	44[b]	(4)	69[b]	(21)	42[b]	(23)	40[b]	(10)	NT	
I	54[b]	(8)	44[b]	(14)	15[b]	(12)	NT		56[b]	(10)
K	79	(14)	74	(7)	58	(15)	NT		NT	
N	48[b]	(14)	50[b]	(18)	60[b]	(21)	NT		NT	
Q	77	(8)	69	(10)	16	(4)	71	(11)	75	(18)
V	98	(4)	94	(8)	90	(4)	98	(4)	85	(14)
X	NT		NT		NT		75[c]	(30)	63[c]	(5)
N-1[d]	8		8		8		5		4	
Mean-1	72		67		40		77		71	
SD-1	20.6		21.7		28.0		23.2		12.9	
CV (%)-1	28.7		32.3		70.6		30.2		18.3	
N-2[e]	5		5		5		4		3	
Mean-2	85		75		40		86		75	
SD-2	9.8		23.2		33.1		12.4		10.5	
CV (%)-2	11.5		30.9		82.6		14.4		13.9	

[a] NT = not tested.

[b] Survival below test acceptable criteria of 70% in WB control sediment at 20 d.

[c] Not included in any mean as WB control sediment was not tested.

[d] N-1, Mean-1, SD1 and CV (%)-1 include all data (except Laboratory X) whether control met acceptable limits or not in WB sediment.

[e] N-2, Mean-2, SD-2 and CV-2 include only data for sediment samples from laboratories that met the 28-d control survival performance acceptability criteria in WB sediment.

Table 17.22 **Interlaboratory Comparison of Dry Weight (Mean mg/Individual ± SD) of *C. tentans* in a Long-term Sediment Exposure Using Five Sediments (WB = West Bearskin, CC = Cole Creek, LS = Little Scioto River, FS = Formulated Sediment (using alpha-cellulose as organic carbon source), and PE = Formulated Sediment (using peat moss as organic carbon source)).**

	Sediment									
Laboratory	WB		CC		LS		FS		PE	
E	1.16	(0.09)	0.71	(0.17)	0.83	(0.32)	1.85	(0.76)	NT[a]	
F	0.94	(0.28)	0.33	(0.07)	3.49	(1.23)	1.84	(0.30)	1.15	(0.19)
H	2.18[b]	(0.13)	0.88[b]	(0.22)	2.85[b]	(0.58)	2.43[b]	(0.30)	NT	
I	1.96[b]	(0.49)	2.00[b]	(0.84)	2.31[b]	(1.17)	NT		2.65	(1.49)
K	1.45	(0.32)	0.71	(0.16)	2.05	(0.29)	NT		NT	
N	1.33[b]	(0.91)	0.99[b]	(0.63)	1.39[b]	(0.66)	NT		NT	
Q	0.79	(0.25)	0.26	(0.04)	1.57	(0.60)	1.13	(0.24)	0.93	(0.45)
V	2.90	(0.73)	1.39	(0.34)	0.66	(0.24)	1.71	(0.52)	2.21	(0.38)
X	NT		NT		NT		1.41[c]	(0.26)	1.83[c]	(0.23)
N-1[d]	8		8		8		5		4	
Mean-1	1.59		0.91		1.89		1.79		1.74	
SD-1	0.71		0.57		0.98		0.46		0.83	
CV (%)-1	44.7		62.6		51.6		25.8		47.7	
N-2[e]	5		5		5		4		3	
Mean-2	1.45		0.68		1.72		1.63		1.43	
SD-2	0.85		0.45		1.14		0.34		0.68	
CV (%)-2	58.6		66.0		66.2		20.9		47.9	

[a] NT = not tested.

[b] Survival below test acceptable criteria of 70% in WB control sediment at 20 d.

[c] Not included in any mean as WB control sediment was not tested.

[d] N-1, Mean-1, SD1 and CV (%)-1 include all data (except Laboratory X) whether control met acceptable limits or not in WB sediment.

[e] N-2, Mean-2, SD-2 and CV-2 include only data for sediment samples from laboratories that met the 28-d control survival performance acceptability criteria in WB sediment.

Note: All dry weight measurements for WB sediment were above the acceptable level of 0.6 mg/organism as dry weight.

Table 17.23 Interlaboratory Comparison of Ash-free Dry Weight (Mean mg/Individual ± SD) of *C. tentans* in a Long-term Sediment Exposure Using Five Sediments (WB = West Bearskin, CC = Cole Creek, LS = Little Scioto River, FS = Formulated Sediment (using alpha-cellulose as organic carbon source), and PE = Formulated Sediment (using peat moss as organic carbon source)).

	Sediment									
Laboratory	WB		CC		LS		FS		PE	
E	0.87	(0.12)	0.54	(0.17)	4.22	(1.80)	1.13	(0.31)	NT[a]	
F	0.65	(0.18)	0.22	(0.03)	2.38	(0.84)	1.18	(0.20)	0.69	(0.19)
H	1.74[b]	(0.13)	0.69[b]	(0.19)	1.93[b]	(0.43)	1.89[b]	(0.40)	NT	
I	NM [a]		NM		NM		NM		NM	
K	1.16	(0.28)	0.51	(0.09)	1.44	(0.29)	NT		NT	
N	0.78[b]	(0.31)	0.99[b]	(0.48)	0.71[b]	(0.47)	NT		NT	
Q	0.57	(0.27)	0.20	(0.03)	1.20	(0.50)	0.83	(0.15)	0.58	(0.26)
V	NM		NM		NM		NM		NM	
X	NT		NT		NT		0.30[c]	(0.04)	0.53[c]	(0.11)
N-1[d]	6		6		6		4		2	
Mean-1	0.96		0.53		1.98		1.26		0.64	
SD-1	0.43		0.30		1.24		0.58		0.08	
CV (%)-1	45.0		56.7		62.6		35.7		12.2	
N-2[e]	4		4		4		3		2	
Mean-2	0.81		0.37		2.31		1.05		0.64	
SD-2	0.43		0.18		1.36		0.19		0.08	
CV (%)-2	53.3		49.6		59.1		18.1		12.7	

[a] NT = not tested; NM = not measured.

[b] Survival below test acceptable criteria of 70% in WB control sediment at 20 d.

[c] Not included in any mean as WB control sediment was not tested.

[d] N-1, Mean-1, SD1 and CV (%)-1 include all data (except Laboratory X) whether control met acceptable limits or not in WB sediment.

[e] N-2, Mean-2, SD-2 and CV-2 include only data for sediment samples from laboratories that met the 28-d control survival performance acceptability criteria in WB sediment.

Note: All dry weight measurements for WB sediment above acceptable level of 0.48 mg/organism as AFDW.

Table 17.24 Interlaboratory Comparison of Percent Emergence (Mean ± SD) of *C. tentans* in a Long-term Sediment Exposure Using Five Sediments (WB = West Bearskin, CC = Cole Creek, LS = Little Scioto River, FS = Formulated Sediment (using alpha-cellulose as organic carbon source), and PE = Formulated Sediment (using peat moss as organic carbon source)).

	Sediment									
Laboratory	WB		CC		LS		FS		PE	
E	65.6	(14.4)	41.7	(19.9)	18.8	(18.8)	75	(21.8)	NT[a]	
F	20.8[b]	(7.7)	5.2[b]	(8.8)	12.5[b]	(16.6)	29.2[b]	(14.1)	31.2[b]	(15.3)
H	28.2[b,c]	(8.9)	28.2[b,c]	(13.3)	46.9[b,c]	(15.4)	26.0[b,c]	(14.4)	NT	
I	11.8[b,c]	(12.0)	22.9[b,c]	(19.2)	5.6[b,c]	(4.1)	NT		8.3[b,c]	(10.7)
K	57.3	(18.6)	24.0	(13.7)	49.0	(10.4)	NT		NT	
N	30.2[b,c]	(17.8)	11.5[b,c]	(6.2)	32.3[b,c]	(10.4)	NT		NT	
Q	56.3	(13.9)	16.7	(10.0)	10.4	(8.6)	26.0	(14.3)	43.8	(20.8)
V	100	(0)	67.7	(16.9)	64.6	(13.2)	NT		67.7	(9.4)
X	NT		NT		NT		46.5[d]	(20.2)	50.7[d]	(24.2)
N-1[e]	8		8		8		4		4	
Mean-1	46.3		27.2		30.0		39.1		37.8	
SD-1	29.1		19.7		21.6		24.0		24.8	
CV (%)-1	62.8		72.4		71.9		61.5		65.7	
N-2[f]	4		4		4		2		2	
Mean-2	69.8		37.5		35.7		50.5		55.8	
SD-2	20.6		22.7		25.4		34.6		16.9	
CV (%)-2	29.5		60.5		71.2		68.6		30.3	

[a] NT = not tested.

[b] Emergence below test acceptable criteria of 50% in WB control sediment.

[c] Survival below test acceptable criteria of 70% in WB control sediment at 20 d.

[d] Not included in any mean as WB control sediment was not tested.

[e] N-1, Mean-1, SD1 and CV (%)-1 include all data (except Laboratory X) whether control met acceptable limits or not in WB sediment.

[f] N-2, Mean-2, SD-2 and CV-2 include only data for sediment samples from laboratories that met the 28-d control survival performance acceptability criteria in WB sediment.

Table 17.25 Interlaboratory Comparison of the Number of Eggs/Female (Mean ± SD) in a Long-term Sediment Exposure Using Five Sediments (WB = West Bearskin, CC = Cole Creek, LS = Little Scioto River, FS = Formulated Sediment (using alpha-cellulose as organic carbon source), and PE = Formulated Sediment (using peat moss as organic carbon source)).

	Sediment									
Laboratory	WB		CC		LS		FS		PE	
E	1258	(429)	523	(124)	1025	(366)	1260	(178)	NT[a]	
F	998	(243)	444	NA	722	(711)	671	(133)	721	(200)
H	1397[b]	(408)	919[b]	(306)	1069[b]	(580)	995[b]	(615)	NT	
I	1261[b]	(225)	538[b]	(117)	NT		NT		988[b]	(290)
K	1023	(177)	538	(117)	835	(86)	NT		NT	
N	1047[b]	(410)	484[b]	(345)	728[b]	(479)	NT		NT	
Q	978	(168)	404	(204)	1190	(126)	1141	(391)	720	(105)
V	1333	(227)	1194	(63)	1127	(191)	NT		1160	(120)
X	NT		NT		NT		828[c]	(286)	827[c]	(214)
N-1[d]	8		8		7		5		5	
Mean-1	1162		631		951		1017		897	
SD-1	168		277		193		255		216	
CV (%)-1	14.4		43.9		20.1		25.1		24.1	
N-2[e]	5		5		5		3		4	
Mean-2	1118		621		980		1024		867	
SD-2	168		325		197		311		254	
CV (%)-2	15.0		52.4		20.1		30.4		29.3	

[a] NT = not tested; NA = not applicable.

[b] Survival below test acceptable criteria of 70% in WB control sediment at 20 d.

[c] Not included in any mean as WB control sediment was not tested.

[d] N-1, Mean-1, SD1 and CV (%)-1 include all data (except Laboratory X) whether control met acceptable limits or not in WB sediment.

[e] N-2, Mean-2, SD-2 and CV-2 include only data for sediment samples from laboratories that met the 28-d control survival performance acceptability criteria in WB sediment.

Note: The number of eggs acceptable criteria (≥800 eggs) was above acceptable level for all laboratories in WB sediment.

Table 17.26 **Interlaboratory Comparison of Percent Hatch (Mean ± SD) of *C. tentans* in a Long-term Sediment Exposure Using Five Sediments (WB = West Bearskin, CC = Cole Creek, LS = Little Scioto River, FS = Formulated Sediment (using alpha-cellulose as organic carbon source), and PE = Formulated Sediment (using peat moss as organic carbon source)).**

	Sediment									
Laboratory	WB		CC		LS		FS		PE	
E	80	(17.0)	37	(33.0)	51	(39.0)	77	(16.1)	NT	
F	99	(0.2)	97	NA	99	NA	97	(2.3)	99	(0.4)
H	93[b]	(3.5)	80[b]	(24.6)	71[b]	(36.5)	74[b]	(49.2)	NT	
I	NM[a]		NM		NM		NM		NM	
K	62[c]	(23.5)	78[c]	(38.5)	74[c]	(14.0)	NT		NT	
N	68[b,c]	(35.8)	47[b,c]	(47.3)	54[b,c]	(40.8)	NT		NT	
Q	80	(35.2)	31	(53.3)	95	(3.2)	89	(19.4)	88	(18.3)
V	91	(8.4)	81	(33.0)	87	(10.8)	NT		96	(1.7)
X	NT		NT		NT		60[d]	(44.0)	80[d]	(27.1)
N-1[e]	7		7		7		4		3	
Mean-1	82		64		76		84		94	
SD-1	13.5		25.6		18.9		10.7		6.0	
CV (%)-1	16.6		39.8		24.9		12.7		6.4	
N-2[f]	4		4		4		3		3	
Mean-2	90		69		94		93		94	
SD-2	9.7		34.3		6.1		5.5		6.0	
CV (%)-2	10.8		49.5		6.5		5.9		6.4	

[a] NT = not tested; NM = not measured; NA = not applicable.

[b] Survival below test acceptable criteria of 70% in WB control sediment at 20 d.

[c] Hatch below test acceptable criteria of 80% in WB control sediment.

[d] Not included in any mean as WB control sediment was not tested.

[e] N-1, Mean-1, SD1 and CV (%)-1 include all data (except Laboratory X) whether control met acceptable limits or not in WB sediment.

[f] N-2, Mean-2, SD-2 and CV-2 include only data for sediment samples from laboratories that met the 28-d control survival performance acceptability criteria in WB sediment.

Section 18
References

Adams, W.J. 1987. Bioavailability of neutral lipophilic organic chemicals contained in sediments. In *Fate and Effects of Sediment-bound Chemicals in Aquatic Systems,* Proceedings of the 6th Pellston Workshop, August 12-17, 1984, Florissant, CO, eds. K.L. Dickson, A.W. Maki, and W.A. Brungs, 219-244. Pergamon Press, New York.

Adams, W.J., Kimerle, R.A., and Mosher, R.G. 1985. An approach for assessing the environmental safety of chemicals sorbed to sediments. In *Aquatic Toxicology and Hazard Evaluation: 7th Symposium*, eds. R.D. Cardwell, R. Purdy, and R.C. Bahner, 429-453. ASTM STP 854. Philadelphia, PA.

Ankley, G.T., Benoit, D.A., Balough, J.C., Reynoldson, T.B., Day, K.E., and Hoke, R.A. 1994a. Evaluation of potential confounding factors in sediment toxicity tests with three freshwater benthic invertebrates. *Environ. Toxicol. Chem.*; 13:627-635.

Ankley, G.T., Benoit, D.A., Hoke, R.A., Leonard, E.N., West, C.W., Phipps, G.L., Mattson, V.R., and Anderson, L.A. 1993. Development and evaluation of test methods for benthic invertebrates and sediments: Effects of flow rate and feeding on water quality and exposure conditions. *Arch. Environ. Contam. Toxicol.* 25:12-19.

Ankley, G.T., Call, D.J., Cox, J.S., Kahl, M.D., Hoke, R.A., and Kosian, P.A. 1994c. Organic carbon partitioning as a basis for predicting the toxicity of chlorpyrifos in sediments. *Environ. Toxicol. Chem.* 13: 621-626.

Ankley, G.T., Collyard, S.A., Monson, P.D., and Kosian, P.A. 1994b. Influence of ultraviolet light on the toxicity of sediments contaminated with polycyclic aromatic hydrocarbons. *Environ. Toxicol. Chem.* 11:1791-1796,

Ankley, G.T., Cook, P.M., Carlson, A.R., Call, D.J., Swenson, J.A., Corcoran, H.F., and Hoke, R.A. 1992b. Bioaccumulation of PCBs from sediments by oligochaetes and fishes: Comparison of laboratory and field studies. *Can. J. Fish. Aquat. Sci.* 49:2080-2085.

Ankley, G.T., Katko, A., and Arthur, J.W. 1990. Identification of ammonia as an important sediment-associated toxicant in the lower Fox River and Green Bay Wisconsin. *Environ. Toxicol. Chem.* 9:313-322.

Ankley, G.T., Lodge, K., Call, D.J., Balcer, M.D., Brooke, L.T., Cook, P.M., Kreis Jr., R.G, Carlson, A.R., Johnson, R.D., Niemi, G.J., Hoke, R.A., West, C.W., Giesy, J.P., Jones, P.D., and Fuying, Z.C. 1992a. Integrated assessment of contaminated sediments in the lower Fox River and Green Bay, Wisconsin. *Ecotoxicol. Environ. Safety* 23:46-63.

Ankley, G.T., Phipps, G.L., Leonard, E.N., Benoit, D.A., Mattson, V.R., Kosian, P.A., Cotter, A.M., Dierkes, J.R., Hansen, D.J., and Mahony, J.D. 1991a. Acid-volatile sulfide as a factor mediating cadmium and nickel bioavailability in contaminated sediment. *Environ. Toxicol. Chem.* 10:1299-1307.

Ankley, G.T. and Schubauer-Berigan, M.K. 1994. Comparison of techniques for the isolation of pore water for sediment toxicity testing. *Arch. Environ. Contam. Toxicol.* 27:507-512.

Ankley, G.T. and Schubauer-Berigan, M.K. 1995. Background and overview of current sediment toxicity identification procedures. *J. Aquat. Ecosystem Health.* 4:133-149.

Ankley, G.T., Schubauer-Berigan, M.K., and Dierkes, J.R. 1991b. Predicting the toxicity of bulk sediments to aquatic organisms using aqueous test fractions: Pore water versus elutriate. *Environ. Toxicol. Chem.* 10:1359-1366.

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

Ankley, G. and Thomas, N. 1992. Interstitial water toxicity identification evaluation approach. In *Sediment Classification Methods Compendium*, pp. 5-1 to 5-14. EPA-823-R-92-006, Washington, DC.

APHA. Part 8010E.4.b. Standard methods for the examination of water and wastewater. 1992. American Public Health Association, Washington, DC. pp. 8-10.

APHA. Standard methods for the examination of water and wastewater. American Public Health Association, Washington, D.C, 1989.

AQUIRE. 1992. Aquatic Toxicity Information Retrieval database and technical support document. USEPA Environmental Research Laboratory, Duluth, MN.

ASTM. 1999a. Standard test methods for measuring the toxicity of sediment-associated contaminants with freshwater invertebrates. E1706-95b. In *Annual Book of ASTM Standards*, Vol. 11.05, Philadelphia, PA.

ASTM. 1999b. Standard guide for collection, storage, characterization, and manipulation of sediments for toxicological testing. E1391-94. In *Annual Book of ASTM Standards,* Vol. 11.05, Philadelphia, PA.

ASTM. 1999c. Standard guide for the determination of bioaccumulation of sediment-associated contaminants by benthic invertebrates. E1688-97a. In *Annual Book of ASTM Standards,* Vol. 11.05, Philidelphia, PA.

ASTM. 1999d. Standard guide for designing biological tests with sediments. E1525-94a. In *Annual Book of ASTM Standards*, Vol. 11.05, Philadelphia, PA.

ASTM. 1999e. Standard guide for conducting acute toxicity tests with fishes, macroinvertebrates, and amphipods. E729-96. In *Annual Book of ASTM Standards*, Vol. 11.05, Philadelphia, PA.

ASTM. 1999f. Standard guide for conducting early life-stage toxicity test with fishes. E1241-98. In *Annual Book of ASTM Standards*, Vol. 11.05, Philadelphia, PA.

ASTM. 1999g. Standard guide for the use of lighting in laboratory testing. E1733-95. In *Annual Book of ASTM Standards*, Vol. 11.05, Philidelphia, PA.

ASTM. 1999h. Standard terminology relating to biological effects and environmental fate. E943-97a. In *Annual Book of ASTM Standards*, Vol. 11.05, Philadelphia, PA.

ASTM. 1999i. Standard practice for conducting bioconcentration tests with fishes and saltwater bivalve mollusks. E1022-94. In *Annual Book of ASTM Standards*, Vol. 11.05, Philadelphia, PA.

Atkinson, A.C. and Donev, A.N. 1992. *Optimum Experimental Designs.* Oxford: Clarendon Press. 328 p.

Bailey, H.C. and Liu, D.H.W. 1980. *Lumbriculus variegatus,* a benthic oligochaete, as a bioassay organism. In *Aquatic Toxicology*, eds. J.C. Eaton, P.R. Parrish, and A.C. Hendricks, 205-215. ASTM STP 707. Philadelphia, PA.

Ball, S.L. and Baker, R.L. 1995. The non-lethal effects of predators and the influence of food availability on the life history of adult *Chironomus tentans* (Diptera: Chironomidae). *Freshwater Biology* 34:1-12.

Bartlett, M.S. 1937. Some examples of statistical methods of research in agriculture and applied biology. *J. Royal Statist. Soc. Suppl.* 4:137-183.

Batac-Catalan, Z. and White, D.S. 1982. Creating and maintaining cultures of *Chironomus tentans* (Diptera: Chironomidae). *Ent. News* 93:54-58.

Becker, D.S., Rose, C.D., and Bigham, G.N. 1995. Comparison of the 10-d freshwater sediment toxicity tests using *Hyalella azteca* and *Chironomus tentans*. *Environ. Toxicol. Chem.* 14:2089-2094.

Beers, Y. 1957. *Introduction to the Theory of Error.* Addison-Wesley Publishing Co. Inc. Reading, MA. p. 26.

Benoit, D.A., Mattson, V.R., and Olson, D.L. 1982. A continuous flow mini-diluter system for toxicity testing. *Water Res.* 16:457-464.

Benoit, D.A., Phipps, G.A., and Ankley, G.T. 1993. A sediment testing intermittent renewal system for the automated renewal of overlying water in toxicity tests with contaminated sediments. *Water Res.* 27:1403-1412.

Benoit, D.A., Sibley, P.K., Juenemann, J.L., and Ankley, G.T. 1997. *Chironomus tentans* life-cycle test: Design and evaluation for use in assessing toxicity of contaminated sediments. *Environ. Toxicol. Chem.* 16:1165-1176.

Besser, J.M., Ingersoll, C.G., Leonard, E., and Mount, D.R. 1998. Effect of zeolite on toxicity of ammonia in freshwater sediments. Implications for sediment toxicity identification evaluation procedures. *Environ. Toxicol. Chem.* 17:2310-2317.

Bligh, E.G. and Dyer, W.J. 1959. A rapid method of total lipid extraction and purification. *Can. J. Biochem. Physiol.* 37:911-917.

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

Borgmann, U. 1994. Chronic toxicity of ammonia to the amphipod *Hyalella azteca*: Importance of ammonium ion and water hardness. *Environ. Pollut.* 86:329-335.

Borgmann, U. 1996. Systematic analysis of aqueous ion requirements of *Hyalella azteca*: A standard artificial medium including the essential bromide ion. *Arch. Environ. Contam. Toxicol.* 30:356-363.

Borgmann, U. and Borgmann, A.I. 1997. The control of ammonia toxicity to *Hyalella azteca* by sodium, potassium, and pH. *Environ. Pollut.* 95:325-331.

Borgmann, U. and Munawar, M. 1989. A new standardized bioassay protocol using the amphipod *Hyalella azteca*. *Hydrobiologia* 188/189:425-531.

Borgmann, U., Ralph, K.M., and Norwood, W.P. 1989. Toxicity test procedures for *Hyalella azteca,* and chronic toxicity of cadmium and pentachlorophenol to *H. azteca, Gammarus fasciatus,* and *Daphnia magna*. *Arch. Environ. Contam. Toxicol.* 18:756-764.

Bovee, E.C. 1949. Studies on the thermal death of *Hyalella azteca* (Saussure). *Biol. Bull.* (Woods Hole) 96:123-128.

Bovee, E.C. 1950. Some effects of temperature on the rates of embryonic, postembryonic, and adult growth in *Hyalella azteca. Proc. Iowa Acad. Sci.* 57:439-444.

Brasher, A.M. and Ogle, R.S. 1993. Comparative toxicity of selenite and selenate to the amphipod *Hyalella azteca*. *Arch. Environ. Contam. Toxicol.* 24:182-186.

Brinkhurst, R.O. 1980. Pollution biology–the North American experience. In *Proceedings of the First International Symposium on Aquatic Oligochaete Biology,* eds. R.O. Brinkhurst and G.C. Cook, 471-475. Plenum Press, New York.

Brinkhurst, R.O. 1986. Guide to the freshwater aquatic microdrile oligochaetes of North America. *Can. Spec. Publ. Fish. Aquatic Sci.* 84. Dept. Fisheries and Oceans, Ottawa, Canada. 259 p.

Brunson, E.L., Ankley, G.T., Burton, G.A., Dwyer, F.J., Ingersoll, C.G., Landrum, P.F., Lee, H. and Phipps, G.L. 1993. Bioaccumulation kinetics and field validation of whole-sediment exposures with the oligochaete, *Lumbriculus variegatus*. Abstract presented at the 14th annual meeting of SETAC, Houston, TX, November 14-18, 1993.

Brunson, E.L., Canfield, T.J., Dwyer, F.J., Kemble, N.E., and Ingersoll, C.G. 1998. Assessment of bioaccumulation from sediments of the upper Mississippi River using field-collected oligochaetes and laboratory-exposed *Lumbriculus variegatus*. *Arch. Environ. Contam. Toxicol.* 35:191-201.

Bryan, G.W. 1976. Some aspects of heavy metal tolerance in aquatic organisms. In *Effects of Pollutants on Aquatic Organisms,* ed. A.P.M. Lockwood, 7-34. Cambridge Univ. Press. New York.

Bufflap, S.T. and Allen, H.E. 1995a. Sediment pore water collection methods for trace metal analysis: A review. *Wat. Res.* 29:165-177.

Bufflap, S.T. and Allen, H.E. 1995b. Comparison of pore water sampling techniques for trace metals. *Wat. Res.* 29:2051-2054.

Bureau of National Affairs, Inc. 1986. U.S. Environmental Protection Agency General Regulation for Hazardous Waste Management. Washington, D.C.

Burton, G.A. 1991. Assessment of freshwater sediment toxicity. *Environ. Toxicol. Chem.* 10:1585-1627.

Burton, G.A., Jr., 1992. Sediment collection and processing factors affecting realism, In *Sediment Toxicity Assessment,* ed. G.A. Burton, Jr., 37-66. Lewis Publ. Boca Raton, FL.

Burton, G.A., Jr. and Ingersoll, C.G. 1994. Evaluating the toxicity of sediments. In *The Assessment of Contaminated Great Lakes Sediment*. U.S. Environmental Protection Agency Report, Region V, EPA/905-B94/002, Chicago, IL.

Burton, G.A., Ingersoll, C.G., Burnett, L.C., Henry, M., Hinman, M.L., Klaine, S.J., Landrum, P.F., Ross, P., and Tuchman, M. 1996a. A comparison of sediment toxicity test methods at three Great Lakes Areas of Concern. *J. Great Lakes Res.* 22:495-511.

Burton, G.A., Nelson, M.K., and Ingersoll, C.G. 1992. Freshwater benthic toxicity tests. In *Sediment Toxicity Assessment,* ed. G.A. Burton, 213-240. Lewis Publishers, Chelsea, MI.

Burton, G.A., Norberg-King, T.J., Ingersoll, C.G., Ankley, G.T., Winger, P.V., Kubitz, J., Lazorchak, J.M., Smith, M.E., Greer, I.E., Dwyer, F.J., Call, D.J., Day, K.E., Kennedy, P., and Stinson, M. 1996b. Interlaboratory study of precision: *Hyalella azteca* and *Chironomus tentans* freshwater sediment toxicity assay. *Environ. Toxicol. Chem*. 15:1335-1343.

Burton, G.A., Stemmer, B.L., Winks, K.L., Ross, P.E., and Burnett, L.C. 1989. A multitrophic level evaluation of sediment toxicity in Waukegan and Indiana Harbors. *Environ. Toxicol. Chem.* 8:1057-1066.

Call, D.J., Balcer, M.D., Brooke, L.T., Lozano, S.J., and Vaishnav, D.D. 1991. Sediment quality evaluation in the lower Fox River and southern Green Bay of Lake Michigan. USEPA Cooperative Agreement Final Report, University of Wisconsin-Superior, Superior, WI.

Call, D.J., Brooke, L.T., Ankley, G.T., Benoit, D.A., and Hoke, R.A. 1994. Appendix G: Biological Effects Testing Procedures. In *Great Lakes Dredged Material Testing and Evaluation Manual*. U.S. Environmental Protection Agency Regions II, III, V; Great Lakes National Program Office; and U.S. Army Corps of Engineers, North Central Division.

Canfield, T.J., Brunson, E.L., Dwyer, F.J., Ingersoll, C.G., and Kemble, N.E. 1998. Assessing sediments from the upper Mississippi River navigational pools using a benthic invertebrate community evaluation and the sediment quality triad. *Arch. Environ. Contam. Toxicol.* 35:202-212.

Canfield, T.J., Dwyer, F.J., Fairchild, J.F., Haverland, P.S., Ingersoll, C.G, Kemble, N.E., Mount, D.R., LaPoint, T.W., Burton, G.A., and Swift, M.C. 1996. Assessing contamination in Great Lakes sediments using benthic invertebrate communities and the sediment quality triad approach. *J. Great Lakes Res.* 22:565-583.

Canfield, T.J., Kemble, N.E., Brumbaugh, W.G., Dwyer, F.J., Ingersoll, C.G., and Fairchild, J.F. 1994. Use of benthic invertebrate community structure and the sediment quality triad to evaluate metal-contaminated sediment in the Upper Clark Fork River, Montana. *Environ. Toxicol. Chem.* 13:1999-2012.

Carlson, A.R., Phipps, G.L., Mattson, V.R., Kosian, P.A., and Cotter, A.M. 1991. The role of acid-volatile sulfide in determining cadmium bioavailability and toxicity in freshwater sediments. *Environ. Toxicol. Chem.* 14:1309-1319.

Carr, R.S. and Chapman, D.C. 1992. Comparison of solid-phase and pore water approaches for assessing the quality of estuarine sediments. *Chemistry and Ecology* 7:19-30.

Carr, R.S. and Chapman, D.C. 1995. Comparison of methods for conducting marine and estuarine sediment pore water toxicity tests, extraction, storage, and handling techniques. *Arch. Environ. Contam. Toxicol.* 28:69-77.

Chapman, P.M., Farrell, M.A., and Brinkhurst, R.O. 1982a. Relative tolerances of selected aquatic oligochaetes to individual pollutants and environmental factors. *Aquat. Toxicol.* 2:47-67.

Chapman, P.M., Farrell, M.A., and Brinkhurst, R.O. 1982b. Relative tolerances of selected aquatic oligochaetes to combinations of pollutants and environmental factors. *Aquat. Toxicol.* 2:69-78.

Chapman, P.M., Power, E.A., and Burton, G.A., Jr. 1992. Integrated assessments in aquatic ecosystems. In *Sediment Toxicity Assessment,* ed. G.A. Burton. Lewis Publishers, Boca Raton, FL.

Chapman, P.M. 1989. Current approaches to developing sediment quality criteria. *Environ. Toxicol. Chem.* 8:589-599.

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

Chekanovskaya, O.V. 1962. Aquatic oligochaeta of the U.S.S.R. Akademiya Nauk SSSR. Moscow, USSR.

Clifford, H. 1991. Aquatic Invertebrates of Canada. The University of Alberta Press, University of Alberta, Edmonton, Alberta, Canada.

Cole and Watkins. 1977. *Hyalella monetzuma*, a new species (Crustacea: Amphipoda) from Montezuma Well, Arizona. *Hydrobiologica* 52:2-3 175-184.

Collyard, S.A., Ankley, G.T., Hoke, R.A., and Goldenstein, T. 1994. Influence of age on the relative sensitivity of *Hyalella azteca* to Diazinon, alkylphenol ethoxylate, copper, cadmium, and zinc. *Arch. Environ. Contam. Toxicol.* 265:110-113.

Connell, D.W., Bowman, M., and Hawker, D.W. 1988. Bioconcentration of chlorinated hydrocarbons from sediment by oligochaetes. *Ecotoxicol. Environ. Safety.* 16:293-302.

Conover, W.J. 1980. *Practical Nonparametric Statistics*. 2nd Ed. John Wiley and Sons, New York, NY, 493 p.

Conover, W.J., Johnson, M.E., and Johnson, M.M. 1981. A comparative study of tests for homogeneity of variances, with applications to the outer continental shelf bidding data. *Technometrics* 23:351-361.

Cook, D.G. 1969. Observations on the life history and ecology of some lumbriculidae (Annelida, Oligochaeta). *Hydrobiologia* 34:561-574.

Cooper, W.E. 1965. Dynamics and production of a natural population of a freshwater amphipod, *Hyalella azteca*. *Ecol. Mong.* 35:377-394.

Craig, G.R. 1984. Bioassessment of sediments: Review of previous methods and recommendations for future test protocols. IEC Beak Consultants, Ltd. Mississauga, Ontario.

Crandall, T., Busack, C.A., and Gall, G.A.E. 1981. An easily constructed recirculating aquarium system for research requiring many small groups of animals. *Aquaculture* 22:193-199.

Curry, L.L. 1962. A survey of environmental requirements for the midge (Diptera: Tendipedidae). In *Biological Problems in Water Pollution*, 3rd seminar, ed. C.M. Tarzwell, 127-141. U.S. Public Health Serv. Publ. 999-WP-25.

Davenport, R. and Spacie, A. 1991. Acute phototoxicity of harbor and tributary sediments from lower Lake Michigan. *J. Great Lakes Res.* 17:51-56.

Davies, R.P. and Dobbs, J.A. 1984. The prediction of bioconcentration in fish. *Wat. Res.* 18:1253-1262.

Davis, R.B., Bailer, A.J., and Oris, J.T. 1998. Effects of organism allocation on toxicity test results. *Environ. Toxicol. Chem.* 17:928-931.

Dawson, T.D., Jenson, J.J., and Norberg-King, T.J. 1999. Laboratory Culture of *Chironomus tentans* for use in toxicity testing: optimum initial egg stocking densities. *Hydrobiologica*: In revision.

Day, K.E., Dutka, B.J., Kwan, K.K., Batista, N., Reynoldson, T.B., and Metcalfe-Smith, J.L. 1995. Correlations between solid-phase microbial screening assays, whole-sediment toxicity tests with macroinvertebrates and *in situ* benthic community structure. *J. Great Lakes Res.* 21:192-206.

Day, K.E., Kirby, R.S., and Reynoldson, T.B. 1994. Sexual dimorphism in *Chironomus riparius* (Meigen): Impact on interpretation of growth in whole-sediment toxicity tests. *Environ. Toxicol. Chem.* 13:35-39.

de Boer, J. 1988. Chlorobiphenyls in bound and nonbound lipids of fishes: Comparison of different extraction methods. *Chemosphere* 17:1803-1810.

DeFoe, D.L. and Ankley, G.T. 1998. Influence of storage time on toxicity of freshwater sediments to benthic macroinvertebrates. *Environ. Pollut.* 99:123-131.

de March, B.G.E. 1977. The effects of photoperiod and temperature on the induction and termination of reproductive resting stage in the freshwater amphipod *Hyalella azteca* (Saussure). *Can. J. Zool.* 55:1595-1600.

de March, B.G.E. 1978. The effects of constant and variable temperatures on the size, growth, and reproduction of *Hyalella azteca* (Saussure). *Can. J. Zool.* 56:1801-1806.

de March, B.G.E. 1981. *Hyalella azteca* (Saussure). In: S.G. Lawrence (ed.), Manual for the culture of selected freshwater invertebrates. *Can. Spec. Pub. Fish. Aquat. Sci.* No. 54, Department of Fisheries and Oceans.

Derr, S.K. and Zabik, M.J. 1972. Biologically active compounds in the aquatic environment: The effect of DDT on the egg viability of *C. tentans*. *Bull. Environ. Contam. Toxicol.* 6:366-369.

DeWitt, T.H., Ditsworth, G.R., and Swartz, R.C. 1988. Effects of natural sediment features on the phoxocephalid amphipod, *Rhepoxynius abronius*: Implications for sediment toxicity bioassays. *Marine Environ. Res.* 25:99-124.

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

DeWoskin, R.S. 1984. *Good Laboratory Practice Regulations: A Comparison.* Research Triangle Institute, Research Triangle Park, NC. 63 p.

Dickson, K.L., Maki, A.W., and Brungs, W.A. 1987. *Fate and Effects of Sediment-bound Chemicals in Aquatic Systems.* Pergamon Press, New York.

Dillon, T.M. and Gibson, A.B. 1986. Bioassessment methodologies for the regulatory testing of freshwater dredged material. Miscellaneous Paper EL-86-6, U.S. Army Engineer Waterways Experiment Station, Vicksburg, MS.

Di Toro, D.M., Mahony, J.H., Hansen, D.J., Scott, K.J., Hicks, M.B., Mayr, S.M., and Redmond, M. 1990. Toxicity of cadmium in sediments: The role of acid-volatile sulfides. *Environ. Toxicol. Chem.* 9:1487-1502.

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

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

Dixon, W.J. and Massey, F.J., Jr. 1983. *Introduction to Statistical Analysis.* 4th Ed. McGraw-Hill Book Company, New York, NY. 678 p.

Driver, E.A. 1977. Chironomid communities in small prairie ponds: Some characteristics and controls. *Freshwater Biol.* 7:121-123.

Driver, E.A., Sugden, L.G., and Kovach, R.J. 1974. Calorific, chemical and physical values of potential duck foods. *Freshwater Biol.* 4:281-29.

Duan, Y., Guttman, S.I., and Oris, J.T. 1997. Genetic differentiation among laboratory populations of *Hyalella azteca*: Implications for toxicology. *Environ. Toxicol. Chem.* 16:691-695.

Duke, B.M., Moore, D.W., and Farrar, J.D. 1996. Effects of preservation on dry weight and length measurements using *Leptocheirus plumulosus*. Abstract presented at the 17th SETAC annual meeting in Washington, DC, November 17-21.

Efron, B. 1982. The Jackknife, the Bootstrap, and other resampling plans. CBMS 38, *Soc. Industr. Appl. Math.,* Philadelphia, PA.

Embody, G.C. 1911. A preliminary study of the distribution, food and reproductive capacity of some freshwater amphipods. *Int. Rev. gesamten Hydrobiol. Biol. Suppl.* 3:1-33.

Enserink, E.L., Kerkhofs, M.J.J., Baltus, C.A.M., and Koeman, J.H. 1995. Influence of food quantity and lead exposure on maturation in *Daphnia magna*: Evidence for a trade-off mechanism. *Functional Ecology* 9:175-185.

Environment Canada. 1997a. Biological Test Method: Test for growth and survival in sediment using the freshwater amphipod *Hyalella azteca.* EPSRN33. Environment Canada, Ottawa, Ontario.

Environment Canada. 1997b. Biological Test Method: Test for growth and survival in sediment using larvae of freshwater midges (*Chironomus tentans* or *Chironomus riparius*). EPSRN32. Environment Canada, Ottawa, Ontario.

Ernsting, G., Zonneveld, C., Isaaks, J.A., and Kroon, A. 1993. Size at maturity and patterns of growth and reproduction in an insect with indeterminate growth. *Oikos* 66:17-26.

Ewell, W.S., Gorsuch, J.W., Kringle, R.O., Robillard, K.A., and Spiegel, R.C. 1986. Simultaneous evaluation of the acute effects of chemicals on seven aquatic species. *Environ. Contam. Toxicol.* 5:831-840.

Fairweather, P.G. 1991. Statistical power and design requirements for environmental monitoring. *Aust. J. Mar. Freshwater Res.* 42:555-567.

Finney, D.J. 1971. *Probit Analysis*. Third edition, Cambridge, University Press, London, 333 p.

Flannagan, J.F. 1971. Toxicity evaluation of trisodium nitrilotriacetate to selected invertebrates and amphibians. *Fish. Res. Board Can. Tech. Rep.* 258. 21 p.

Folch, J., Lees, M., and Stanley, G.H.S. 1957. A simple method for isolation and purification of total lipids from animal tissue. *J. Biol. Chem.* 226:497-509.

Food and Drug Administration. 1978. Good laboratory practices for nonclinical laboratory studies. Part 58. Fed. Reg. 43(247): 60013-60020 (December 22, 1978).

France, R.L. 1992. Biogeographical variation in size-specific fecundity of the amphipod *Hyalella azteca*. *Crustaceana* 62:240-248.

Gardner, W.S., Frez, W.A., Cichocki, E.A., and Parrish, C.C. 1985. Micromethods for lipids in aquatic invertebrates. *Limnol. Oceanog.* 30:1099-1105.

Gaston, G.R., Bartlett, J.H., McAllister, A.P., and Heard, R.W. 1996. Biomass variations of estuarine macrobenthos preserved in ethanol and formalin. *Estuaries* 19:674-679.

Gauss, J.D., Woods, P.E., Winner, R.W., and Skillings J.H. 1985. Acute toxicity of copper to three life stages of *Chironomus tentans* as affected by water hardness-alkalinity. *Environ. Poll.* (Ser. A) 37:149-157.

Geisler, F.S. 1944. Studies on the post-embryonic development of *Hyalella azteca* (Saussure). *Biol. Bull.* 86:6-22.

Giesy, J.P., Graney, R.L., Newsted, J.L., Rosiu, C.J., Benda, A., Kreis, Jr., R.G., and Horvath, F.J. 1988. Comparison of three sediment bioassay methods using Detroit River sediments. *Environ. Toxicol. Chem.* 7:483-498.

Gill, J.L. 1978. *Design and Analysis of Experiments in the Animal and Medical Sciences.* Vol. 3. Appendices. The Iowa State University Press, Ames, IA, 173 p.

Gobas, F.A.P.C., Bedard, D.C., Ciborowski, J.J.H., and Haffner, G.D. 1989. Bioaccumulation of chlorinated hydrocarbons by the mayfly (*Hexagenia limbata*) in Lake St. Clair. *J. Great Lakes Res.* 15:581-588.

Green, R.H. 1979. *Sampling Design and Statistical Methods for Environmental Biologists*. Wiley-Interscience. New York. 257 p.

Greer, I.E. 1993. Standard operating procedures for culture of chironomids (SOP B5.25 dated 02/18/93) and *Hyalella azteca* (SOP B5.38 dated 09/17/93). USGS, Columbia, MO.

Grothe, D.R. and Kimerle, R.A. 1985. Inter- and Intralaboratory variability in *Daphnia magna* effluent toxicity test results. *Environ. Toxicol. Chem.* 4:189-192.

Hall, W.S., Patoczka, J.B., Mirenda, R.J., Porter, B.A., and Miller, E. 1989. Acute toxicity of industrial surfactants to *Mysidopsis bahia*. *Arch. Environ. Contam. Toxicol.* 18:765-772,

Hamilton, M.A., Russo, R.C., and Thurston, R.V. 1977. Trimmed Spearman-Karber method for estimating median lethal concentrations in toxicity bioassays. *Environ. Sci. Technol*. 11:714-719.

Hanes, E.C., Ciborowski, J.J.H., and Corkum, L.D. 1991. Standardized rearing materials and procedures for *Hexagenia,* a benthic aquatic bioassay organism. Annual report submitted to the Research Advisory Committee, Ontario Ministry of the Environment, Toronto, Ontario, September 1991.

Hargrave, B.T. 1970a. The utilization of benthic microflora by *Hyalella azteca. J. Animal Ecology* 39:427-437.

Hargrave, B.T. 1970b. Distribution, growth, and seasonal abundance of *Hyalella azteca* (amphipod) in relation to sediment microflora. *J. Fish. Res. Bd. Can.* 27:685-699.

Harkey, G.A., Kane Driscoll, S., and Landrum, P. 1997. Effect of feeding in 30-day bioaccumulation assays using *Hyalella azteca* in fluoranthene-dosed sediment. *Environ. Toxicol. Chem.* 16:762-769.

Harkey, G.A., Landrum, P.F., and Klaine, S.J. 1994. Preliminary studies on the effect of feeding during whole-sediment bioassays using *Chironomus riparius* larvae. *Chemosphere* 28:597-606.

Harrahy, E.A. and Clements, W.H. 1997. Toxicity and bioaccumulation of a mixture of heavy metals in *Chironomus tentans* (Diptera: Chironomidae) in synthetic sediment. *Environ. Toxicol. Chem.* 16:317-327.

Herbes, S.E. and Allen, C.P. 1983. Lipid quantification of freshwater invertebrates: Method modification for microquantification. *Can. J. Fish. Aquat. Sci.* 40:1315-1317.

Hilsenhoff, W.L. 1966. The biology of *Chironomus plumulosus* (Diptera: Chironomidae) in Lake Winnebago, Wisconsin. *Ann. Entomol. Soc. Am.* 59: 365-473.

Hoke, R.A., Ankley, G.T., Cotter, A.M., Kosian, P.A., Phipps, G.L., and Vanderrneiden, F.M. 1994. Evaluation of equilibrium partitioning theory for predicting acute toxicity of field-collected sediments contaminated with DDT, DDE and DDD to the amphipod *Hyalella azteca. Environ. Toxicol. Chem.* 13:157-166.

Hoke, R.A., Giesy, J.P., Ankley, G.T., Newsted, J.L., and Adams, R.J. 1990. Toxicity of sediments from western Lake Erie and the Maumee River at Toledo, Ohio, 1987: Implications for current dredged material disposal practices. *J. Great Lakes Res.* 16:457-470.

Hoke, R.A., Kosian, P.A., Ankley, G.T., Cotter, A.M., Vandenneiden, F.M., Phipps, G.L., and Durhan, E.J. 1995. Check studies with *Hyalella azteca* and *Chironomus tentans* in support of the development of a sediment quality criterion for dieldrin. *Environ. Toxicol. Chem.* 14:435-443.

Hornig, C.E. 1980. Use of the aquatic oligochaete, *Lumbriculus variegatus,* for effluent biomonitoring. EPA-600/D-80-005. National Technical Information Service, Springfield, VA.

Horwitz, A.J., Elrick, K.A., and Colberg, M.R. 1992. The effect of membrane filtration artifacts on dissolved trace element concentrations. *Wat. Res.* 26:753-763.

Hurlbert, S.H. 1984. Pseudoreplication and the design of ecological field experiments. *Ecol. Mono.* 54:187-211.

Ingersoll, C.G. 1995. Sediment toxicity tests. In *Fundamentals of Aquatic Toxicology,* Second Edition, ed. G. Rand, 231-255. Taylor and Francis, Washington, DC.

Ingersoll, C.G., Ankley, G.T., Benoit, D.A., Burton, G.A., Dwyer, F.J., Greer, I.E., Norberg-King, T.J., and Winger, P.V. 1995. Toxicity and bioaccumulation of sediment-associated contaminants with freshwater invertebrates: A review of methods and applications. *Environ. Toxicol. Chem.* 14:1885-1894.

Ingersoll, C.G., Brunson, E.L., Dwyer, F.J., Hardesty, D.K., and Kemble, N.E. 1998. Use of sublethal endpoints in sediment toxicity tests with the amphipod *Hyalella azteca*. *Environ. Toxicol. Chem.* 17:1508-1523.

Ingersoll, C.G., Buckler, D.R., Crecelius, E.A., and La Point, T.W. 1993. U.S. Fish and Wildlife Service and Battelle final report for the USEPA GLNPO assessment and remediation of contaminated sediment (ARCS) project: Biological assessment of contaminated Great Lakes sediment. EPA-905-R93-006, Chicago, IL.

Ingersoll, C.G., Dillon, T., and Biddinger, R.G. (eds.). 1997. Methodological uncertainty in sediment ecological risk assessment. In *Ecological Risk Assessments of Contaminated Sediment*, 389 p. SETAC Press, Pensacola, FL.

Ingersoll, C.G., Dwyer, F.J., Burch, S.A., Nelson, M.K., Buckler, D.R., and Hunn, J.B. 1992. The use of freshwater and saltwater animals to distinguish between the toxic effects of salinity and contaminants in irrigation drainwater. *Environ. Toxicol. Chem.* 11:503-511.

Ingersoll C.G., Dwyer F.J., and May, T.W. 1990. Toxicity of inorganic and organic selenium to *Daphnia magna* (Cladocera) and *Chironomus riparius* (Diptera). *Environ. Toxicol. Chem.* 9:1171-1181.

Ingersoll, C.G., Haverland, P.S., Brunson, E.L., Canfield, T.J., Dwyer, F.J., Henke, C.E., and Kemble, N.E. 1996. Calculation and evaluation of sediment effect concentrations for the amphipod *Hyalella azteca* and the midge *Chironomus riparius*. *J. Great Lakes Res.* 22:602-623.

Ingersoll, C.G. and Nelson, M.K. 1990. Testing sediment toxicity with *Hyalella azteca* (Amphipoda) and *Chironomus riparius* (Diptera). In *Aquatic Toxicology and Risk Assessment*, 13th volume, eds. W.G. Landis and W.H. van der Schalie, 93-109. ASTM STP 1096. Philadelphia, PA.

Jones, R.A. and Lee, G.F. 1988. Toxicity of U.S. waterway sediments with particular reference to the New York Harbor area. In *Chemical and Biological Characterization of Sludges, Sediments, Dredge Spoils, and Drilling Muds,* eds. J.J. Lichtenburg, F.A. Winter, C.I. Weber, and L. Fradkin, 403-417. STP 976. American Society for Testing and Materials, Philadelphia, PA.

Kane Driscoll, S., Landrum, P.F., and Tigue, E. 1997. Accumulation and toxicokinetics of fluoranthene in water-only exposures with freshwater amphipods. *Environ. Toxicol Chem.* 16:754-761.

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

Kates, M. 1986. *Techniques of Lipidology (Isolation, Analysis, and Identification of Lipids),* 2nd rev. ed. Elsevier Science Pub. Co., New York. 464 p.

Kemble, N.E., Brumbaugh, W.G., Brunson, E.L., Dwyer, F.J., Ingersoll, C.G., Monda, D.P., and Woodward, D.F. 1994. Toxicity of metal-contaminated sediments from the Upper Clark Fork River, MT, to aquatic invertebrates in laboratory exposures. *Environ. Toxicol. Chem.* 13:1985-1997.

Kemble, N.E., Dwyer, F.J., Ingersoll, C.G., Dawson, T.D., and Norberg-King, T.J. 1999. Tolerance of freshwater test organisms to formulated sediments for use as control materials in whole-sediment toxicity tests. *Environ. Toxicol. Chem.* 18: 222-230.

Kemble, N.E., Brunson,E.L., Canfield, T.J., Dwyer, F.J., and Ingersoll, C.G. 1998. Assessing sediment toxicity from navigational pools of the upper Mississippi River using a 28-d *Hyalella azteca* test. *Arch. Environ. Contam. Toxicol.* 35: 181-190.

Kemp, P.F. and Swartz, R.C. 1988. Acute toxicity of interstitial and particle-bound cadmium to a marine infaunal amphipod. *Marine Environ. Res.* 26:135-153.

Kielty, T.J., White, D.S., and Landrum, P.F. 1988a. Short-term lethality and sediment avoidance assays with endrin-contaminated sediment and two oligochaetes from Lake Michigan. *Arch. Environ. Contam. Toxicol.* 17:95-101.

Kielty, T.J., White, D.S., and Landrum, P.F. 1988b. Sublethal responses to endrin in sediment by *Limnodrilus hoffmeisteri* (Tubificidae), and in mixed culture with *Stylodrilus heringianus* (Lumbriculidae). *Aquat. Toxicol.* 13:227-250.

Knezovich, J.P., Harrison, F.L., and Wilhelm, R.G. 1987. The bioavailability of sediment-sorbed organic chemicals: A review. *Water Air Soil Pollut.* 32:233-245.

Kubitz, J.A., Besser, J.M., and Giesy, J.P. 1996. A two-step experimental design for a sediment bioassay using growth of amphipod *Hyalella azteca* for the test endpoint. *Environ. Toxicol. Chem.* 15:1783-1792.

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

Kukkonen, J. and Landrum, P.F. 1995. Effects of sediment-bound polydimethylsiloxane on the bioavailability and distribution of benzo[a]pyrene in lake sediment to *Lumbriculus variegatus*. *Environ. Toxicol. Chem.* 14:523-531.

Lacey, R., Watzin, M.C., and McIntosh, A.W. 1999. Sediment organic matter content as a confounding factor in toxicity tests with *Chironomus tentans. Environ. Toxicol. Chem.* 18: 231-236.

Lake, J.L., Rubinstein, N.I., Lee II, H., Lake, C.A., Heltshe, J., and Pavignano, S. 1990. Equilibrium partitioning and bioaccumulation of sediment-associated contaminants by infaunal organisms. *Environ. Toxicol. Chem.* 9:1095-1106.

Lamberson, J.O. and Swartz, R.C. 1988. Use of bioassays in determining the toxicity of sediment to benthic organisms. In *Toxic Contaminants and Ecosystem Health: A Great Lakes Focus,* ed. M.S. Evans, 257-279. John Wiley and Sons, New York.

Lamberson J.O. and Swartz, R.C. 1992. Spiked-sediment toxicity test approach. In *Sediment Classification Methods Compendium*, pp. 4-1 to 4-10. EPA-823-R-92-006, Washington, DC.

Landrum, P.F. 1989. Bioavailability and toxicokinetics of polycyclic aromatic hydrocarbons sorbed to sediments for the amphipod *Pontoporeia hoyi*. *Environ. Sci. Technol.* 23:588-595.

Landrum, P.F. and Faust, W.R. 1992. Variation in the bioavailability of polycyclic aromatic hydrocarbons sorbed to sediments for the amphipod *Pontoporeia hoyi. Environ. Toxicol. Chem.* 11:1197-1208.

Landrum, P.F. and Scavia, D. 1983. Influence of sediment on anthracene uptake, depuration, and biotransformation by the amphipod, *Hyalella azteca*. *Can. J. Fish. Aquat. Sci.* 40:298-305.

Landrum, P.F., Faust, W.R., and Eadie, B.J. 1989. Bioavailability and toxicity of a mixture of sediment-associated chlorinated hydrocarbons to the amphipod *Pontoporeia hoyi*. In *Aquatic Toxicology and Hazard Assessment,* eds. U.M. Cowgill and L.R. Williams, 315-329. ASTM STP 1027. Philadelphia, PA.

Lauritsen, D.D., Mozley, S.C., and White, D.S. 1985. Distribution of oligochaetes in Lake Michigan and comments on their use as indices of pollution. *J. Great Lakes Res.* 11:67-76.

Lee, D.R. 1980. Reference toxicants in quality control of aquatic bioassays. In *Aquatic Invertebrate Bioassays,* eds. A.L. Buikema and J. Cairns, Jr., 188-199. ASTM STP 715, Philadelphia, PA.

Lee, II, H., Boese, B.L., Pelletier, J., Randall, R.C., and Specht, D.T. 1990. Method to estimate gut uptake efficiencies for hydrophobic organic pollutants. *Environ. Toxicol. Chem.* 9:215-220.

Leppanen, C.J. and Maier, K.J. 1998. An inexpensive and efficient modular water-renewal system for bulk sediment toxicity testing. *Environ. Toxicol. Chem.* 17: 969-971.

Liber, K., Call, D.J., Dawson, T.D., Whiteman, F.W., and Dillon, T.M. 1996. Effects of *Chironomus tentans* larval growth retardation on adult emergence and ovipositing success: Implications for interpreting freshwater sediment bioassays. *Hydrobiologia* 323:155-167.

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

Long, E.R. and Morgan, L.G. 1990. The potential for biological effects of sediment-sorbed contaminants tested in the national status and trends program. NOAA Technical Memorandum NOS OMA 52, Seattle, WA.

Luoma, S.N. and Bryan, G.W. 1982. A statistical study of environmental factors controlling concentrations of heavy metals in the burrowing bivalve *Scrobicularia plana* and the polychaete *Nereis diversicolor*. *Estuarine Coastal Shelf Sci.* 15:95-108.

MacDonald, D.D., Carr, R.S., Calder, F.D., Long, E.R., and Ingersoll, C.G. 1996. Development and evaluation of sediment quality guidelines for Florida coastal waters. *Ecotoxicology* 5:253-278.

Maki, A.W. 1977. Modifications of continuous flow test methods for small aquatic organisms. *Prog. Fish. Cult.* 39:172-174.

Marcus, A.H. and Holtzman, A.P. 1988. A robust statistical method for estimating effects concentrations in short-term fathead minnow toxicity tests. Manuscript *submitted* to the Criteria and Standards Division, U. S. Environmental Protection Agency, by Battelle Washington Environmental Program Office, Washington, DC, June 1988, under EPA Contract No. 69-03-3534. 39 p.

Mayer, Jr., F.L. and Ellersieck, M.R. 1986. Manual of acute toxicity: Interpretation and data base for 410 chemicals and 66 species of freshwater animals. U.S. Fish and Wildlife Service Resource Publication 160, Washington, DC.

McLarney, W.O., Henderson, S., and Sherman, M.S. 1974. A new method for culturing *Chironomus tentans* Fabricius larvae using burlap substrate in fertilized pools. *Aquaculture* 4:267-276.

McNulty, E.W., Dwyer, F.J., Ellersieck, M.R., Greer, I.E., Ingersoll, C.G., and Rabeni, C.F. 1999. Evaluation of the ability of reference-toxicity tests to identify stress in laboratory populations of the amphipod *Hyalella azteca*. *Environ. Toxicol. Chem*. 18: 544-548.

Merritt, R.W. and Cummins, K.W. 1996. *An Introduction to the Aquatic Insects of North America.* 3rd edition, Kent Hunt Publishing Co., Dubuque, Iowa, 862 p.

Milani, D., Day, K.E., McLeay, D.J., and Kirby, R.S. 1996. Recent intra- and interlaboratory studies related to the development and standardization of Environment Canada's biological test methods for measuring sediment toxicity using freshwater amphipods (*Hyalella azteca*) or midge larvae (*Chironomus riparius*). Environment Canada, Burlington, ONT.

Mood, A.M., Graybill, F.A., and Boes, D.C. 1984. *Introduction to the Theory of Statistics*, 3rd ed. McGraw-Hill Book Company. New York. 546 p.

Moore, D.W. and Dillon, T.M. 1993. The relationship between growth and reproduction in the marine polychaete *Nereis* (*Neanthes*) *arenaceodentata* (Moore): Implications for chronic sublethal sediment bioassays. *J. Exp. Mar. Biol. Ecol.* 173:231-246.

Moore, D.W. and Farrar, J.D. 1996. Effect of growth on reproduction in the freshwater amphipod *Hyalella azteca*. *Hydrobiologia* 328:127-134.

Moore, D.W., Dillon, T.M., and Gamble, E.W. 1996. Long-term storage of sediments: Implications for sediment toxicity testing. *Environ. Pollut.* 89: 341-342.

Mount, D.I. and Brungs, W.A. 1967. A simplified dosing apparatus for fish toxicology studies. *Water Res.* 1:21-30.

Mount, D.R., Dawson, T.D., and Burkhard, L.P. 1999. Implications of gut purging for tissue residues determined in bioaccumulation testing of sediments with *Lumbriculus variegatus. Environ. Toxicol. Chem.* 18:1244-1249.

National Research Council (NRC). 1989. *Contaminated Marine Sediments—Assessment and Remediation.* NRC, National Academy Press, Washington, DC.

Naylor, C. 1993. Guide to the preparation of artificial sediment for use in tests with *Chironomus riparius.* Standard operating procedure. Department of Animal and Plant Sciences, University of Sheffield, United Kingdom.

Nebeker, A.V. and Miller, C.E. 1988. Use of the amphipod crustacean *Hyalella azteca* in freshwater and estuarine sediment toxicity tests. *Environ. Toxicol. Chem.* 7:1027-1033.

Nebeker, A.V., Cairns, M.A., Gakstatter, J.H., Malueg, K.W., Schuytema, G.S., and Krawczyk, D.F. 1984a. Biological methods for determining toxicity of contaminated freshwater sediments to invertebrates. *Environ. Toxicol. Chem.* 3:617-630.

Nebeker, A.V., Cairns, M.A., and Wise, C.M. 1984b. Relative sensitivity of *Chironomus tentans* life stages to copper. *Environ. Toxicol. Chem.* 3:151-158.

Nebeker, A.V., Griffis, W.L., Wise, C.M., Hopkins, E., and Barbitta, J.A. 1989. Survival, reproduction and bioconcentration in invertebrates and fish exposed to hexachlorobenzene. *Environ. Toxicol. Chem.* 8:601-611.

Nebeker, A.V., Onjukka, S.T., and Cairns, M.A. 1988. Chronic effects of contaminated sediment on *Daphnia magna* and *Chironomus tentans. Bull. Environ. Contam. Toxicol*. 41:574-581.

Nebeker, A.V., Onjukka, S.T., Stevens, D.G., Chapman, G.A., and Dominguez, S.E. 1992. Effects of low dissolved oxygen on survival, growth and reproduction of *Daphnia, Hyalella* and *Gammarus. Environ. Toxicol. Chem.* 11:373-379.

New, M.B., Scholl, J.P., McCarty, J.C., and Bennett, J.P. 1974. A recirculating system for experimental aquaria. *Aquaculture* 3:95-103.

Newman, M.C. 1995. Quantitative Methods in Aquatic Toxicology. In *Advances in Trace Sustances Research,196-200.* Lewis Publisher, Boca Raton, FL.

Obana, H., Hori, S., and Kushimoto, T. 1981. Determination of polycyclic aromatic hydrocarbons, i.e., marine samples, by high performance liquid-chromatography. *Bull. Environ. Contam. Toxicol.* 26:613-620.

Oliver, B.G. 1984. Uptake of chlorinated organics from anthropogenically contaminated sediments by oligochaete worms. *Can. J. Fish. Aquat. Sci.* 41:878-883.

Oliver, B.G. 1987. Biouptake of chlorinated hydrocarbons from laboratory-spiked and field sediments by oligochaete worms. *Environ. Sci. Technol.* 21:785-790.

Oliver, B.J. and Niimi, A.J. 1983. Bioconcentration of chlorobenzenes from water by rainbow trout: Correlations with partition coefficients and environmental residues. *Environ. Sci. Technol.* 17:287-291.

Oliver, D.R. 1971. Life history of the chironomidae. *Ann. Rev. Entomol.* 16:211-230.

Oris, J.T. and Giesy, J.P. 1985. The photoenhanced toxicity of anthracene to juvenile sunfish (*Lepomis* spp.). *Aquat. Toxicol.* 6:133-146.

Pascoe, D., Williams, K.A., and Green, D.W.J. 1989. Chronic toxicity of cadmium to *Chironomus riparius* Meighen–effects upon larval development and adult emergence. *Hydrobiologia* 175: 109-115.

Pennak, R.W. 1989. *Freshwater Invertebrates of the United States.* John Wiley and Sons, Inc., New York. 628 p.

Pennak, R.W. and Rosine, W.A. 1976. Distribution and ecology of Amphipoda (Crustacea) in Colorado. *Am. Midl. Nat.* 96:325-331.

Pesch, C. and Morgan, D. 1987. Influence of sediment in copper toxicity tests with the polychaete *Neanthes arenaceodentata. Wat. Res.* 12:747-751.

Phipps, G.L., Ankley, G.T., Benoit, D.A., and Mattson, V.R. 1993. Use of the aquatic oligochaete *Lumbriculus variegatus* for assessing the toxicity and bioaccumulation of sediment-associated contaminants. *Environ. Toxicol. Chem.* 12:269-274.

Phipps, G.L., Mattson, V.R., and Ankley, G.T. 1995. The relative sensitivity of three benthic test species to 10 chemicals. *Arch. Environ. Toxicol. Chem.* 28:281-286.

Pittinger, C.A. and Adams, W.J. 1997. Assessing ecological risks to benthic species in product and technology development. In *Ecological Risk Assessment of Contaminated Sediment,* eds. C.G. Ingersoll, T. Dillon, G.R. Biddinger, 11-21. SETAC Press, Pensacola, FL.

Plumb, Jr., R.H. 1981. Procedures for handling and chemical analysis of sediment and water samples. Technical committee on criteria for dredged and fill material. USEPA-USACE. EPA-4805572010. USEPA Great Lakes Laboratory. Grosse Ile, MI.

Postma, J.F., van Kleunen, A., and Admiraal, W. 1995. Alterations in life-history traits of *Chironomus riparius* (Diptera) obtained from metal contaminated rivers. *Arch. Environ. Contam. Toxicol.* 29: 469-475.

Ramirez-Romero, P. 1997. Evaluation of the acute and chronic toxicity of fluoranthene-spiked sediments in the presence of UV light using the amphipod, *Hyalella azteca*. Ph.D. Dissertation, Maimi University, Oxford, OH, 130 pages.

Randall, R.C., Lee II, H., Ozretich, R.J., Lake, L.J., and Pruell, R.J. 1991. Evaluation of selected lipid methods for normalizing pollutant bioaccumulation. *Environ. Toxicol. Chem.* 10:1431-1436.

Rees, M. and Crawley, M.J. 1989. Growth, reproduction and population dynamics. *Functional Ecology* 3:645-653.

Reish, D.J. 1988. The use of toxicity testing in marine environmental research. In *Marine Organisms as Indicators,* eds. D.F. Soule and G.S. Kleppel, Chapter 10, 213-245. Springer-Verlag, New York.

Reynoldson, T.B., Day, K.E., Clarke, C., and Milani, D. 1994. Effect of indigenous animals on chronic endpoints in freshwater sediment toxicity tests. *Environ. Contam. Toxicol.* 13:973-977.

Roach, R.W., Carr, R.S., Howard, C.L., and Cain, D.W. 1992. Assessment of produced water impacts in Galveston Bay System. U.S. Fish and Wildlife Report, Clear Lake Ecological Services Office, Houston, TX.

Robbins, J.A., Kielty, T.J., White, D.S., and Edgington, D.N. 1989. Relationships among tubificid abundances, sediment composition and accumulation rates in Lake Erie. *Can. J. Fish. Aquat. Sci.* 46:223-231.

Roberts, J.B., deFrietas, A.W.S, and Gidney, M.A.J. 1977. Influence of lipid pool size on bioaccumulation of the insecticide chlordane by northern redhorse suckers *Moxostoma macrolepidotum. J. Fish. Res. Board Can.* 34:89-97.

Rohlf, F.J. and Sokal, R.R. 1981. *Statistical Tables.* W.H. Freeman and Company, New York, NY.

Rottmann, R.W. and Campton, D.E. 1989. Multiple-tank aquarium system with recirculating water for laboratory studies of freshwater fishes. *Prog. Fish-Cult.* 51:238-243.

Rubinstein, N.I., Lake, J.L., Pruell, R.J., Lee II, H., Taplin, B., Heltshe, J., Bowen, R., and Pavignano, S. 1987. Predicting bioaccumulation of sediment-associated organic contaminants: Development of a regulatory tool for dredged material evaluation. Internal report. U.S. EPA-600/x-87/368, Narragansett, RI. 54 p.+ appendices.

Sadler, W.O. 1935. Biology of the midge *Chironomus tentans* Fabricius, and methods for its propagation. Cornell Univ. Agric. Exp. Station Mem. 173. 25 p.

Sarda, N. and Burton, G.A., Jr. 1995. Ammonia variation in sediments: Spatial, temporal and method-related effects. *Environ. Toxicol. Chem.* 14: 1499-1506.

Satterthwaite, F.W. 1946. An approximate distribution of estimates of variance components. *Biom. Bull.* 2:110-114.

Schaeffer, D.J. and Janardan, K.G. 1978. Theoretical comparison of grab and composite sampling programs. *Biom. J.* 20:215-227.

Schlekat, C.E., McGee, B.L., Boward, D.M., Reinharz, E., Velinsky, D.J., and Wade, T.L. 1994. Tidal river sediments in the Washington, D.C. area. III. Biological effects associated with sediment contamination. *Estuaries* 17:334-344.

Schmitt, C.J. and Finger, S.E. 1987. The effects of sample preparation of measured concentrations of eight elements in edible tissues of fish for streams contaminated by lead mining. *Arch. Environ. Contam. Toxicol.* 16:185-207.

Schmitt, C.J., Zajicek, J.L., and Peterman, P.H. 1990. National contaminant biomonitoring program: Residues of organochlorine chemicals in U.S. freshwater fish, 1976-1984. *Arch. Environ. Contam. Toxicol.* 19:748-781.

Schubauer-Berigan, M.K. and Ankley, G.T. 1991. The contribution of ammonia, metals, and nonpolar organic compounds to the toxicity of sediment interstitial water from an Illinois River tributary. *Environ. Toxicol. Chem.* 10:925-939.

Schubauer-Berigan, M.K., Dierkes, J.R., Monson, P.D., and Ankley, G.T. 1993. The pH-dependent toxicity of Cd, Cu, Ni, Pb and Zn to *Ceriodaphnia dubia, Pimephales promelas, Hyalella azteca,* and *Lumbriculus variegatus. Environ. Toxicol. Chem.* 12:1261-1266.

Schubauer-Berigan, M.K., Monson, P.D., West, C.W., and Ankley , G.T. 1995. Influence of pH on the toxicity of ammonia to *Chironomus tentans* and *Lumbriculus variegatus*. *Environ. Toxicol. Chem.* 14:713-717.

Schults, D.W., Ferraro, S.P., Smith, L.M., Roberts, F.A., and Poindexter, C.K. 1992. A comparison of methods for collecting interstitial water for trace organic compounds and metal analyses. *Wat. Res.* 26:989-995.

Schuytema, G.S., Krawczyk, D.F., Griffis, W.L., Nebeker, A.V., Robideaux, M.L., Brownawell, B.J., and Westall, J.C. 1988. Comparative uptake of hexachlorobenzene by fathead minnows, amphipods and oligochaete worms from water and sediment. *Environ. Toxicol. Chem.* 7:1035-1044.

Schuytema, G.S., Nebeker, A.V., Griffis, W.L., and Miller, C.E. 1989. Effects of freezing on toxicity of sediments contaminated with DDT and Endrin. *Environ. Toxicol. Chem.* 8:883-891.

Scott, K.J. 1989. Effects of contaminated sediments on marine benthic biota and communities. In *Contaminated Marine Sediments—Assessment and Remediation.* pp. 132-154. National Research Council. National Academy Press, Washington, DC.

Shapiro, S.S. and Wilk, M.B. 1965. An analysis of variance test for normality (complete samples). *Biometrika* 52:591-611.

Shuba, P.J., Tatem, H.E., and Carroll, J.H. 1978. Biological assessment methods to predict the impact of open-water disposal of dredged material. U.S. Army Corps of Engineers Technical Report D-78-5Q, Washington, DC.

Sibley, P.K., Ankley, G.T., Cotter, A.M., and Lenoard, E.N. 1996. Predicting chronic toxicity of sediments spiked with zinc: An evaluation of the acid-volatile sulfide (AVS) model using a life-cycle test with the midge *Chironomus tentans*. *Environ. Toxicol. Chem.* 15:2102-2112.

Sibley, P.K., Benoit, D.A., and Ankley, G.T. 1997a. The significance of growth in *Chironomus tentans* sediment toxicity tests: Relationship to reproduction and demographic endpoints. *Environ. Toxicol. Chem.* 16:336-345.

Sibley, P.K., Monson, P.D., and Ankley, G.T. 1997b. The effect of gut contents on dry weight estimates of *Chironomus tentans* larvae: Implications for interpreting toxicity in freshwater sediment toxicity tests. *Environ. Toxicol. Chem.* 16:1721-1726.

Sibley, P.K., Benoit, D.A., and Ankley, G.T. 1998. Life cycle and behavioural assessments of the influence of substrate particle size on *Chironomus tentans* (Diptera: Chironomidae) in laboratory assays. *Hydrobiologia* 361:1-9.

Siegfried, W.R. 1973. Summer food and feeding of the ruddy duck in Manitoba. *Can. J. Zool.* 51:1293-1297.

Sijm, R.T.H., Haller, M., and Schrap, S.M. 1997. Influence of storage on sediment characteristics and of drying sediment on sorption coefficients of organic contaminants. *Bull. Environ. Contam. Toxicol.* 58:961-968.

Smith, M.E., Lazorchak, J.M., Herrin, L.E., Brewer-Swartz, S., and Thoney, W.T. 1997. A reformulated, reconstituted water for testing the freshwater amphipod, *Hyalella azteca*. *Environ. Toxicol. Chem.* 16:1229-1233.

Smock, L.A. 1980. Relationships between body size and bio-mass of aquatic insects. *Freshwater Biol.* 10:375-383.

Smrcheck, J.C. and Zeeman, M. 1998. Assessing risks to ecological systems from chemicals. Chapter 3 In: *Handbook for Environmental Risk Assessment and Management*, ed. P. Calow, pp. 24-90. Blackwell Science Ltd., London.

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

Southerland, E., Kravitz, M., and Wall, T. 1992. Management framework for contaminated sediments (the U.S. EPA sediment management strategy). In *Sediment Toxicity Assessment,* ed. G.A. Burton, Jr., pp. 341-370. Lewis Publishers, Chelsea, MI.

Snedecor, G.W. and Cochran, G.C. 1989. *Statistical Methods*. 8th edition, 507 p. The Iowa State University Press, Ames, IA.

Spacie, A. and Hamelink, J.L. 1982. Alternative models for describing the bioconcentration of organics in fish. *Environ. Toxicol. Chem.* 1:309-320.

Spencer, D.R. 1980. The aquatic oligochaetes of the St. Lawrence Great Lakes region. In *Proceedings of the First International Symposium on Aquatic Oligochaete Biology*, eds. R.O. Brinkhurst, and D.G. Cook, pp. 115-164. Plenum Press, New York.

Sprague, J.B. 1963. Resistance of four freshwater crustaceans to lethal high temperature and low oxygen. *J. Fish. Res. Board Can.* 20:387-415.

Steel, R.G.D., and Torrie, J.A. 1980. *Principles and Procedures of Statistics*. McGraw-Hill Book Co., New York.

Steevens, J.A. and Benson, W.H. 1998. *Hyalella azteca* 10-day sediment toxicity test: Comparison of growth measurement endpoints. *Environ. Toxicol. Water Qual.* 13:243-248.

Stehly, G.R., Landrum, P.F., Henry, M.G., and Klemm, C. 1990. Toxicokinetics of PAHs in *Hexagenia*. *Environ. Toxicol. Chem.* 9:167-174.

Stemmer, B.L., Burton, G.A., Jr., and Sasson-Brickson, G. 1990a. Effect of sediment spatial variance and collection method on cladoceran toxicity and indigenous microbial activity determinations. *Environ. Toxicol. Chem.* 9:1035-1044.

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

Stephan, C.E., Mount, D.I., Hansen, D.J., Gentile, J.H., Chapman, G.A., and Brungs, W.A. 1985. Guidelines for deriving numerical national water quality criteria for the protection of aquatic organisms and their uses. PB85-227049, National Technical Information Service, Springfield, VA.

Strong, D.R. 1972. Life history variation among populations of an amphipod (*Hyalella azteca*). *Ecology* 53:1103-1111.

Suedel, B.C. and Rodgers, Jr., J.H. 1994. Development of formulated reference sediments for freshwater and estuarine sediment toxicity testing. *Environ. Toxicol. Chem.* 13:1163-1176.

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

Swartz, R.C. 1989. Marine sediment toxicity tests. In *Contaminated Marine Sediments—Assessment and Remediation.* National Research Council, National Academy Press, Washington, DC. pp. 115-129.

Swartz, R.C., Cole, F.A., Lamberson, J.O., Ferraro, S.P., Schults, D.W., DeBen, W.A., Lee II, H., and Ozretich, R.J. 1994. Sediment toxicity, contamination, and amphipod abundance at a DDT and dieldrin-contaminated site in San Francisco Bay. *Environ. Toxicol. Chem.* 13:949-962,

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

Taylor, J.K. 1987. Quality assurance of chemical measurements. Lewis Publ., Inc., Chelsea, MI.

Tetra Tech, Inc. 1986. ODES statistical power analysis. Draft report, prepared for Office of Marine and Estuarine Protection, USEPA Contract NO. 68-01-6938, TC-3953-03. Bellevue, WA.

Tomasovic, M.J., Dwyer, F.J., Greer, I.E., and Ingersoll, C.G. 1995. Recovery of known-age *Hyalella azteca* (Amphipoda) from sediment toxicity tests. *Environ. Toxicol. Chem.* 14:1177-1180.

Tomasovic, M., Dwyer, F.J., and Ingersoll, C.G. 1994. Recovery of *Hyalella azteca* from sediment. *Environ. Toxicol. Chem.* 13:1311-1314.

Topping, M.S. 1971. Ecology of larvae of *Chironomus tentans* (Diptera: Chironomidae) in saline lakes in central British Columbia. *Can. Entomol.* 193:328-338.

Townsend, B.E., Lawrence, S.G., and Flannagan, J.F. 1981. *Chironomus tentans* Fabricius. In *Manual for the Culture of Selected Freshwater Invertebrates,* ed. S.G. Lawrence, pp. 109-126. Can. Spec. Publ. Fish Aquatic Sci., no. 54. Dept. of Fisheries and Oceans, Winnipeg, Canada.

Unger, P.D., Hague, K., and Schwartz, A.L. 1993. Clinical comparison of a formaldehyde free histological fixative (NoTox) to neutral buffered formalin. *International Academy of Pathologists,* March 1993, New Orleans, LA.

USEPA. 1973. Biological field and laboratory methods for measuring the quality of surface waters and effluents. EPA-670/4-73/001, Cincinnati, OH.

USEPA. 1978. The *Selenastrum capricornutum* Printz algal assay bottle test. EPA-600/9-78-018, Corvallis, OR.

USEPA. 1979a. Handbook for analytical quality assurance in water and wastewater laboratories. EPA-600/4-79-019, Cincinnati, OH.

USEPA. 1979b. Methods for chemical analysis of water and wastes. EPA-600/4-79-020, Cincinnati, OH.

USEPA. 1980a. Proposed good laboratory practice guidelines for toxicity testing. Paragraph 163.60-6. Fed. Reg. 45:26377-26382 (April 18, 1980).

USEPA. 1980b. Physical, chemical, persistence, and ecological effects testing; good laboratory practice standards (proposed rule). 40 CFR 772, Fed. Reg. 45:77353-77365 (November 21, 1980).

USEPA. 1985. Methods for measuring the acute toxicity of effluents to freshwater and marine organisms. 3rd Ed. EPA-600/4-85/013, Cincinnati, OH.

USEPA. 1986a. Quality criteria for water. EPA-440/5-86-001, Washington, D.C.

USEPA. 1986b. Occupational health and safety manual. Office of Administration, EPA, Washington, DC.

USEPA. 1989a. Evaluation of the apparent effects threshold (AET) approach for assessing sediment quality. Report of the sediment criteria subcommittee. SAB-EETFC-89-027, Washington, DC.

USEPA. 1989b. Guidance manual: Bedded sediment bioacumulation tests. EPA-600/X-89/302, USEPA, ERL-N, Pacific Ecosystems Branch, Newport, OR.

USEPA. 1989c. Short-term methods for estimating the chronic toxicity of effluents and receiving waters to freshwater organisms. EPA-600/4-89/001, Cincinnati, OH.

USEPA. 1989d. Toxic Substance Control Act (TSCA); Good laboratory practice standards, Final Rule, Federal Register 54: 34034-34050. August 17.

USEPA. 1990a. Evaluation of the equilibrium partitioning (EqP) approach for assessing sediment quality. Report of the sediment criteria subcommittee of the ecological processes and effects committee. EPA-SAB-EPEC-90-006, Washington, DC.

USEPA. 1990b. Evaluation of the sediment classification methods compendium. Report of the sediment criteria subcommittee of the ecological processes and effects committee. EPA-SAB-EPEC-90-018, Washington, DC.

USEPA. 1990c. Analytical procedures and quality assurance plan for determination of xenobiotic chemical contaminants in fish. EPA-600/3-90/023, Duluth, MN.

USEPA. 1990d. Analytical procedures and quality assurance plan for determination of PCDD/PCDF in fish. EPA-600/3-90/022, Duluth, MN.

USEPA. 1991a. Methods for measuring the acute toxicity of effluents and receiving waters to freshwater and marine organisms. Fourth edition. EPA-600/4-90/027F, Cincinnati, OH.

USEPA. 1991b. Sediment toxicity identification evaluation: Phase I (Characterization), Phase II (Identification) and Phase III (Confirmation). Modifications of effluent procedures. EPA-600/6-91/007, Duluth, MN.

USEPA. 1991c. Technical support document for water quality-based toxic control. EPA-505/2-90/001, Washington, DC.

USEPA. 1992a. Proceedings from workshop on tiered testing issues for freshwater and marine sediments. Washington, DC, September 16-18.

USEPA. 1992b. An SAB report: Review of sediment criteria development methodology for nonionic organic contaminants. Report of the sediment criteria subcommittee of the ecological processes and effects committee. EPA-SAB-EPEC-93-002, Washington, DC.

USEPA. 1992c. Sediment classification methods compendium. EPA-823-R-92-006, Washington, DC.

USEPA. 1993a. Standard operating procedures for culturing *Hyalella azteca* (ERL-D-SOP-CTI-016), *Chironomus tentans* (ERL-D-SOP-CTI-015), and *Lumbriculus variegatus* (ERL-D-SOP-CTI-017). USEPA, Environmental Research Laboratory, Duluth, MN.

USEPA. 1993b. Test methods for evaluating solid waste, physical/chemical methods (SW-846), Third edition. Washington, D.C. 20460.

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

USEPA. 1994b. Methods for measuring the toxicity of sediment-associated contaminants with estuarine and marine amphipods. EPA-600/R-94/025. Narragansett, RI.

USEPA. 1994c. Short-term methods for estimating the chronic toxicity of effluents and receiving waters to freshwater organisms. Third edition. EPA-600/4-91/002, Cincinnati, OH.

USEPA. 1994d. Short-term methods for estimating the chronic toxicity of effluents and receiving waters to marine and estuarine organisms. Second edition. EPA-600/4-91/003, Cincinnati, OH.

USEPA. 1994e. EPA requirements for quality assurance project plans for environmental data operations. Draft interim final. EPA QA/R-5. Office of Research and Development, Washington, DC.

USEPA. 1995. QA/QC guidance for sampling and analysis of sediments, water, and tissues for dredged material evaluations—chemical evaluations. EPA 823-B-95-001. Office of Water, Washington, DC.

USEPA. 1996a. Calculation and evaluation of sediment effect concentrations for the amphipod *Hyalella azteca* and the midge *Chironomus riparius*. EPA 905-R96-008, Chicago, IL.

USEPA. 1996b. Marine Toxicity Identification Evaluation (TIE). Phase I Guidance Document. EPA/600/R-96/054. Eds. R.M. Burgess, K.T. Ho, G.E. Morrison, G. Chapman, D.L. Denton. U.S. Environmental Protection Agency, National Health and Environmental Effects Research Laboratory, Narragansett, Rhode Island.

USEPA. 1997a. The incidence and severity of sediment contamination in surface waters of the United States, Volume 1: National sediment quality survey. EPA 823-R-97-006. Office of Science and Technology, Washington, DC.

USEPA. 1997b. The incidence and severity of sediment contamination in surface waters of the United States, Volume 2: Data summaries for areas of probable concern. EPA 823-R-97-007. Office of Science and Technology, Washington, DC.

USEPA. 1997c. The incidence and severity of sediment contamination in surface waters of the United States, Volume 3: National sediment contaminant point source inventory. EPA 823-R-97-008. Office of Science and Technology, Washington, DC.

USEPA. 1998. Contaminated sediment management strategy. EPA 823-R-98-001. Office of Water, Washington, DC.

USEPA. 1999. Methods for collection, storage, and manipulation of sediments for chemical and toxicological analyses. (May 14, 1999 draft).

USEPA-USACE (U.S. Army Corps of Engineers). 1977. Ecological evaluation of proposed discharge of dredged material into ocean waters. Technical committee on criteria for dredged and fill material. Environmental Effects Laboratory. U.S. Army Engineer Waterways Experiment Station. Vicksburg, MS.

USEPA-USACE. 1991. Evaluation of dredged material proposed for ocean disposal testing manual. EPA-503/8-91/001, Washington, DC.

USEPA-USACE. 1998a. Evaluation of dredged material proposed for discharge in waters of the U.S.- testing manual. EPA-823-B-98-004, Washington, DC.

USEPA-USACE. 1998b. Method for assessing the chronic toxicity of marine and estuarine sediment-associated contaminants with the amphipod *Leptocheirus plumulosus.* Office of Research and Development, Newport, OR. (August, 1998 draft).

Van Rees, K.C.J., Sudlicky, E.A., Suresh, P., Rao, C., and Reddy, K.R. 1991. Evaluation of laboratory techniques for measuring diffusion coefficients in sediments. *Environ. Sci. Technol.* 25:1605-1611.

Vassilaro, D.L., Stoker, P.W., Booth, G.M., and Lee, M.L. 1982. Capillary gas chromatographic determination of polycyclic aromatic compounds in vertebrate tissue. *Anal. Chem.* 54:106-112.

Wall, V.D., London, J., Warren, J.E., Gossett, R., Wenholz, M.D., and Klaine, S.J. 1998. Development of a continuous-flow renewal system for sediment toxicity testing. *Environ. Toxicol. Chem.* 17: 1159-1164.

Walsh, G.E., Weber, D.E., Simon, T.L., and Brashers, L.K. 1991. Toxicity tests of effluents with marsh plants in water and sediment. *Environ. Toxicol. Chem.* 10:517-525.

Walters, D.B. and Jameson, C.W. 1984. *Health and Safety for Toxicity Testing*. Butterworth Publications, Woburn, MA.

Wen, Y.H. 1993. Sexual dimorphism and male choice in *Hyalella azteca*. *Am. Midl. Nat.*129:153-160.

Wentsel, R., McIntosh, A., and Atchison, G. 1977a. Sublethal effects of heavy metal contaminated sediment on midge larvae (*Chironomus tentans*). *Hydrobiologia.* 56:53-156.

Wentsel, R., McIntosh, A., and Anderson, V. 1977b. Sediment contamination and benthic invertebrate distribution in a metal-impacted lake. *Environ. Pollut*. 14:187-193.

Wentsel, R., McIntosh, A., and McCafferty, P.C. 1978. Emergence of the midge *Chironomus tentans* when exposed to heavy metal contaminated sediments. *Hydrobiologia* 57:195-196.

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

West, C.W., Phipps, G.L., Hoke, R.A., Goldenstein, T.A., Vandertneiden, F.M., Kosian, P.A., and Ankley, G.T. 1994. Sediment core vs grab sample: Evaluation of contamination and toxicity at a DDT contaminated site. *Ecotox Env. Safety* 208-220.

Whiteman, F.W., Ankley, G.T., Dahl, M.D., Rau, D.M., and Balcer, M.D. 1996. Evaluation of interstitial water as a route of exposure to ammonia in sediment tests with macroinvertebrates. *Environ. Toxicol. Chem.* 15:794-801.

Wiederholm, T., Wiederholm, A.M., and Goran, M. 1987. Freshwater oligochaetes. *Water Air Soil Pollut.* 36:131-154.

Williams, K.A., Green, D.W.J., Pascoe, D., and Gower, D.E. 1986b. The acute toxicity of cadmium to different larval stages of *Chironomus riparius* (Diptera:Chironomidae) and its ecological significance for pollution regulation. *Oecologia* 70:362-366.

Williams, K.A., Green, D.J., Pascoe, D., and Gower, D.E. 1987. Effect of cadmium on oviposition and egg viability in *Chironomus riparius* (Diptera: Chironomidae). *Bull. Environ. Contam. Toxicol.* 38: 86-90.

Williams, L.R. 1993. Methods for the EPA's regulatory program. *Environmental Testing and Analysis* 2:36.

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

Winer, B.J. 1971. *Statistical Principles in Experimental Design.* 2nd Ed. McGraw-Hill Book Company, New York, NY. 907 p.

Winger, P.V. and Lasier, P.J. 1993. Vacuum-extracted pore water toxicity testing. In *Environmental Toxicology and Risk Assessment,* 2nd volume, eds. J.W. Gorsuch, F.J. Dwyer, C.G. Ingersoll, T.W. La Point, 640-662. ASTM STP 1173. Philadelphia, PA.

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

Word, J.Q., Ward, J.A., Franklin, L.M., Cullinan, V.I., and Kiesser, S.L. 1987. Evaluation of equilibrium partitioning theory for estimating the toxicity of the nonpolar organic compound DDT to the sediment dwelling amphipod *Rhepoxynius abronius*. USEPA Criteria and Standards Division. SCD #11, Washington, DC.

Zar, J.H. 1984. *Biostatistical Analysis*. 2nd Ed. Prentice-Hall, Inc., Englewood Cliffs, NJ. 717 p.

Zumwalt, D.C., Dwyer, F.J., Greer, I.E., and Ingersoll, C.G. 1994. A water-renewal system that accurately delivers small volumes of water to exposure chamber. *Environ. Toxicol. Chem.* 1311-1314.

Appendix A
Exposure Systems

A.1 Renewal of overlying water is recommended during sediment tests (Section 11.3, 12.3, 13.3, 14.3, 15.3). The overlying water can be replaced manually (e.g., siphoning) or automatically. Automatic systems require more equipment and initially take more time to build, but manual addition of water takes more time during a test. In addition, automated systems generally result in less suspension of sediment compared to manual renewal of water.

A.2 At any particular time during the test, flow rates through any two test chambers should not differ by more than 10%. Mount and Brungs (1967) diluters have been modified for sediment testing, and other diluter systems have also been used (Maki, 1977; Ingersoll and Nelson, 1990; Benoit et al., 1993; Zumwalt et al., 1994; Brunson et al., 1998; Wall et al., 1998; Leppanen and Maier, 1998). The water-delivery system should be calibrated before a test is started to verify that the system is functioning properly. Renewal of overlying water is started on Day -1 before the addition of test organisms or food on Day 0. Water-delivery systems are described by Benoit et al. (1993) in Section A.3 and by Zumwalt et al. (1994) in Section A.4. A 60-mL syringe with a mesh screen over the end can be used to manually remove and replace overlying water (J. Lazorchak, USEPA, Cincinnati, OH, personal communication).

A.3 Benoit et al. (1993) describe a sediment testing intermittent-renewal (STIR) system (stationary or portable) for invertebrate toxicity testing with sediment. The STIR system has been used to conduct both short-term and long-term sediment toxicity tests with amphipods and midges (Sections 11, 12, 14, 15). Either stationary or portable systems enable the maintenance of acceptable water quality (e.g., dissolved oxygen) by automatically renewing overlying water in sediment tests at rates ranging from 1 to 21 volume renewals/d. The STIR system not only reduces the labor associated with renewal of overlying water but also affords a gentle exchange of water that results in virtually no sediment suspension. Both gravity-operated systems can be installed in a compact vented enclosure. The STIR system has been used for conducting 10-d whole-sediment tests with *Chironomus tentans, Hyalella azteca* and *Lumbriculus variegatus.*

A.3.1 STIR systems described in Benoit et al. (1982) can be modified to conduct sediment tests and at the same time maintain their original capacity to deliver varying concentrations of toxicants for water-only toxicity tests. A STIR system (stationary or portable) solely for sediment toxicity tests was designed, which offers a simple, inexpensive approach for the automated renewal of variable amounts of overlying water (Figures A.1 and A.2). This system is described below. The system can be built as a two-unit system (Section A.3.2) or with more exposure treatments (Section A.3.4). All exposure systems consist of exposure holding tanks, head tanks, head tank support stands, and a water bath (Section A.3.2 and A.3.3). The automated delivery system includes design descriptions for a support stand, water renewal supply, and water-delivery apparatus (Section A.3.4).

A.3.2 Two-unit Portable STIR System Construction (Figures A.1 and A.2)

A.3.2.1 Exposure Holding Tanks *(2)* (Figure A.3).

1. Outer diameter: 15.8 cm wide x 29.3 cm long x 11.7 cm high

2. Cutting dimensions: (double-strength glass, 3 mm)

2 Bottoms:	15.8 cm x 29.3 cm
4 Sides:	11.4 cm x 28.7 cm
4 Ends:	11.4 cm x 15.8 cm

3. Hole: 1.6 cm centered between sides and 7.2 cm from bottom edge of 11.4 cm high end piece.

4. Standpipe Height: 10.3 cm above inside of tank bottom.

A.3.2.2 Head Tanks *(2)* (4-L capacity; Figure A.3)

1. Outer diameter: 15.8 cm wide x 24 cm long x 14.5 cm high

2. Cutting dimensions: (acrylic plastic, 6 mm)

2 Bottoms:	15.8 cm x 24 cm
4 Sides:	13.9 cm x 22.8 cm
4 Ends:	13.9 cm x 15.8 cm

3. Acrylic plastic sheets should be cut with a smooth cutting fine toothed table saw blade. Dimension cut pieces can most easily be glued together with Weld-On® #16 clear-thickened cement for acrylic

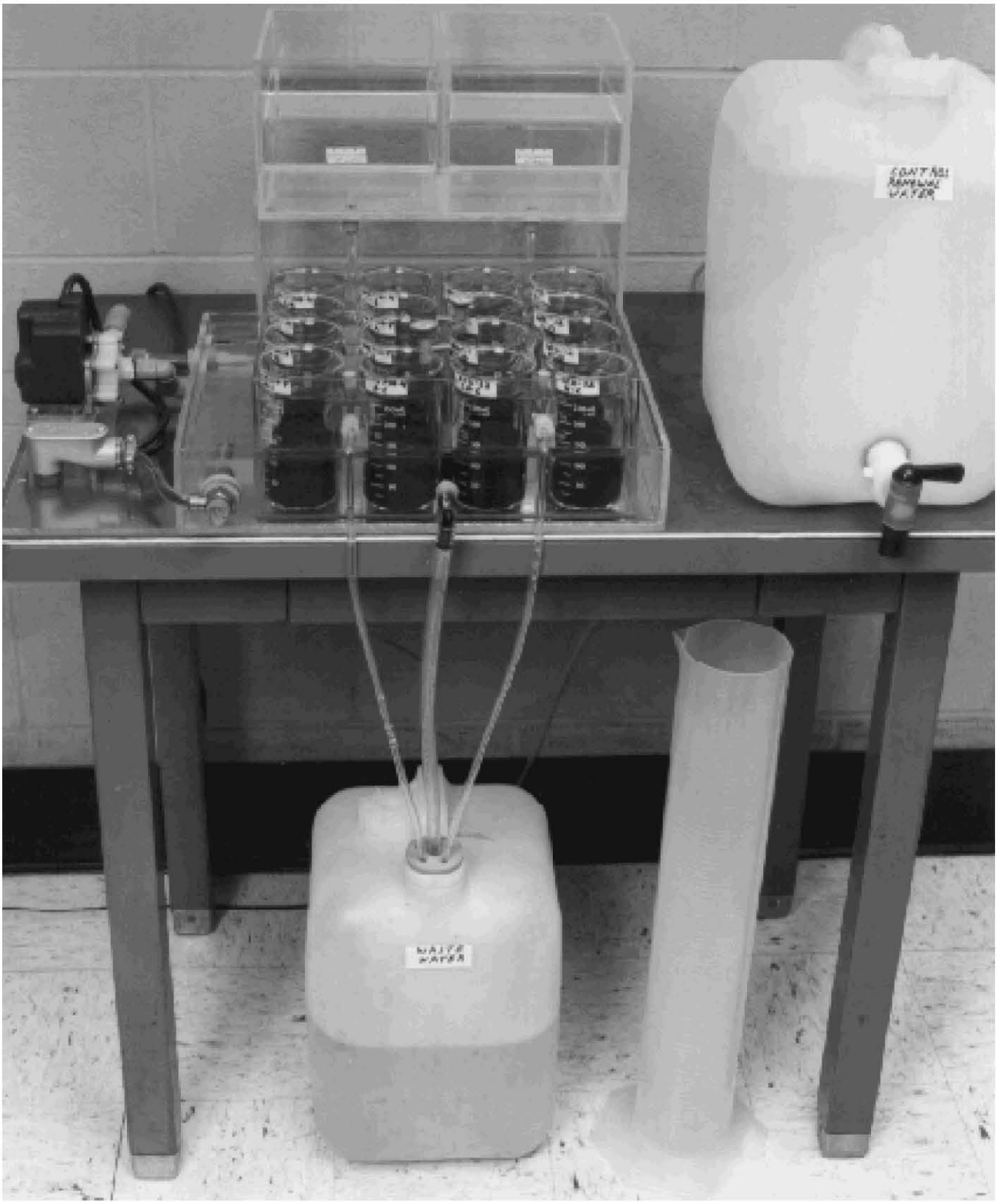


Figure A.1 Portable table top STIR system described in Benoit et al. (1993).

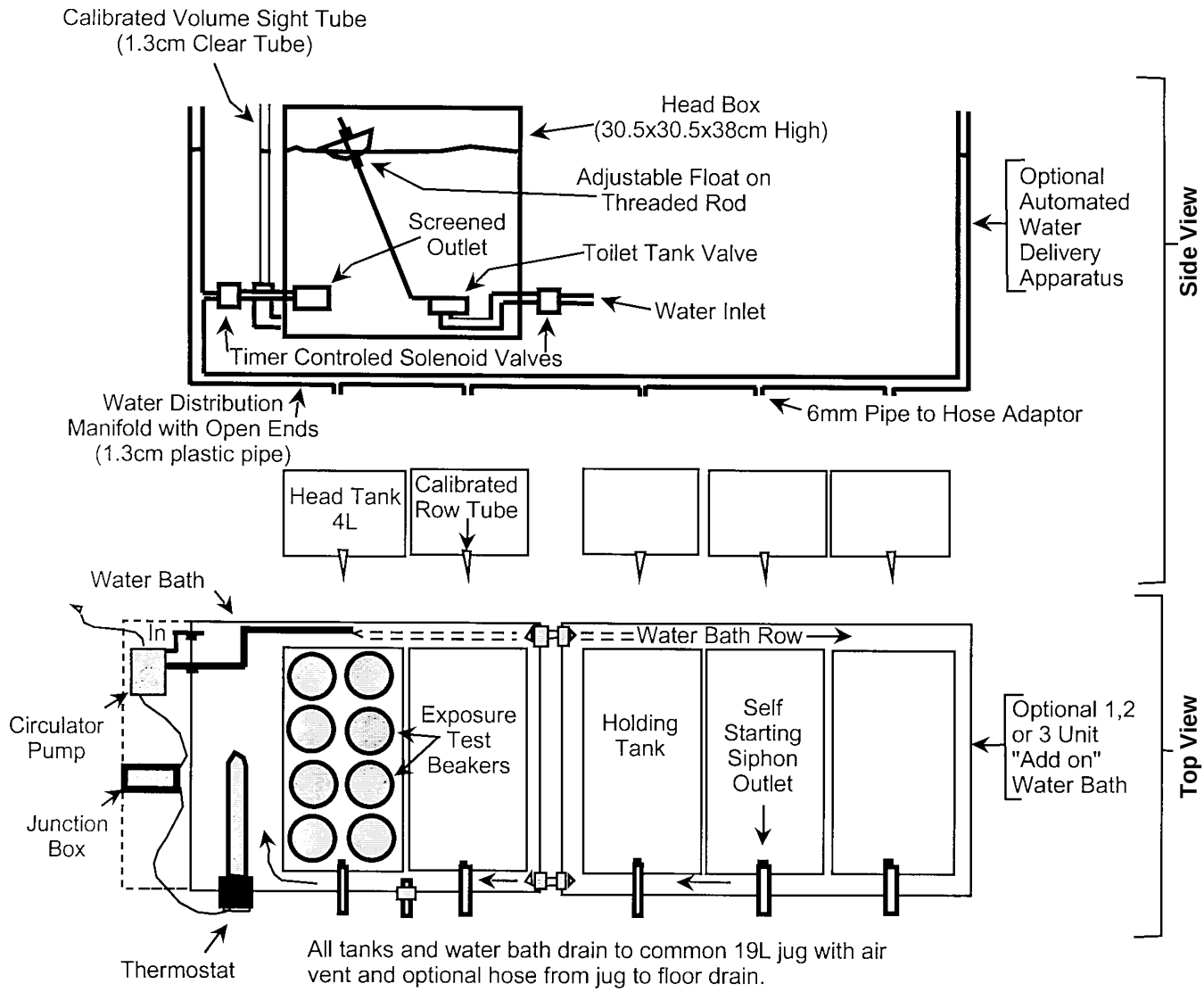


Figure A.2 Portable table top STIR system with several additional options as described in Benoit et al. (1993).

plastic (Industrial Polychemical Service, P.O. Box 471, Gardena, CA, 90247).

4. Hole: 1.6 cm centered between sides and 2 cm from front edge of 24-cm-long bottom piece. Holes can most easily be drilled in acrylic plastic by using a wood spade bit and drill press.

5. Flow Tubes: 10-mL pipet tip initially cut off at the 6-mL mark and inserted flush with top of #0 stopper. Top of stopper should be inserted nearly flush with head tank bottom. With 2 L of water in head tank, calibrate flow tube to deliver 32 mL/min.

A.3.2.3 Head Tank Support Stand *(1) (Figure A.3)*

1. Outer diameter: 16.7 cm wide x 33.7 cm long x 17.8 cm high

2. Cutting dimensions: (acrylic plastic, 6 mm)

1 Bottom:	16.7 cm x 33.7 cm
2 Sides:	17.2 cm x 32.5 cm
2 Ends:	17.2 cm x 16.7 cm

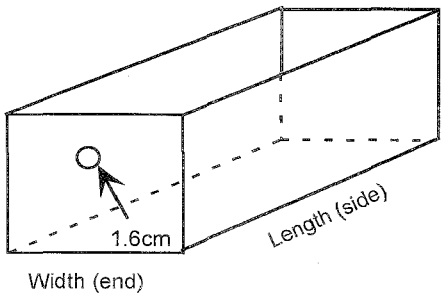
3. Size is such that both head tanks fit into support stand for storage and transport.

A.3.2.4 Water Bath *(1) (Figure A.3)*

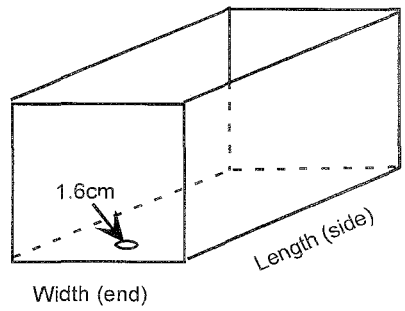
1. Outer diameter: 33 cm wide x 40.6 cm long x 7.4 cm high

2. Cutting dimensions: (acrylic plastic, 6 mm)

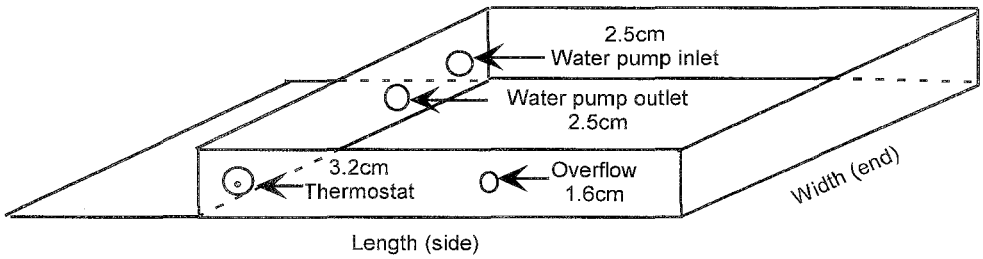
1 Bottom:	33 cm x 55.9 cm
2 Ends:	33 cm x 6.8 cm
2 Sides:	39.4 cm x 6.8 cm



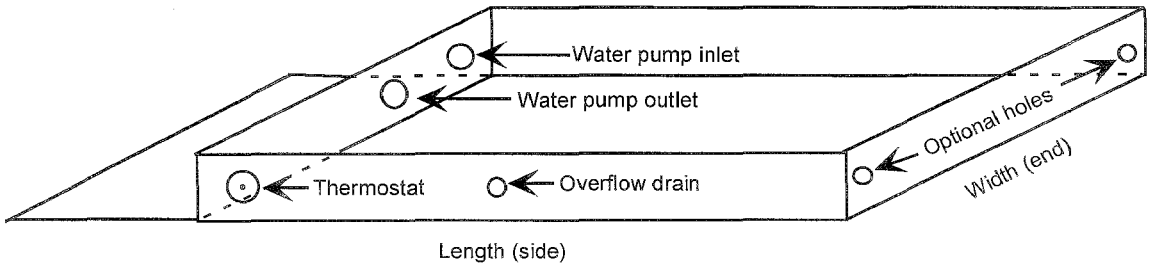
Exposure Holding Tank



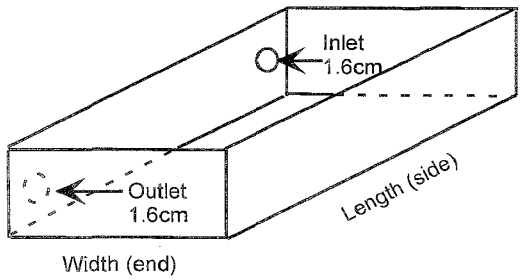
Head Tank



Basic Water Bath



Basic Water Bath with Optional Holes for Water Bath



Add-on Water Bath for One Additional Unit

Figure A.3 Tanks for the STIR system in Benoit et al. (1993).

3. Holes:
   a. Overflow drain; 1.6 cm centered 2.9 cm from bottom edge of 39.4-cm-long side piece and 17.8 cm from right edge.
   b. Thermostat; 3.2 cm centered 2.5 cm from bottom edge of 39.4-cm-long side piece and 3.2 cm from left edge.
   c. Water pump outlet; 2.5 cm centered 2.5 cm from bottom edge of 33-cm-long end piece and 8.3 cm from back edge.
   d. Water pump inlet; 2.5 cm centered 2.5 cm from bottom edge of 33-cm-long end piece and 2.0 cm from back edge.

4. A small 90° elbow made of glass or plastic is attached to the water pump inlet tube and turned downward so the circulator pump will not pick up air at the water surface.

5. The bottom piece for the water bath includes 15.3-cm extension for motor mount and the thermostat electrical junction box.

6. Motor Mount: 5.1 cm wide x 11.4 cm long x 3.8 cm thick mount made from 6 pieces of 6-mm acrylic plastic. Four of these pieces are glued together. The other two pieces are glued together, motor attached to the edge with two screws and the two pieces (with motor attached) are then screwed to the top of the four pieces. The entire unit is then glued to water bath extension after 6-mm PVC piping is attached and secured with stoppers to the inlet and outlet water bath holes.

7. Thermostat Conduit Junction Box: (1.3-cm small left back (SLB)) is attached to the water bath extension by screwing a 1.3-cm PVC plug into junction box and securing this plug with a screw, countersunk up through the bottom and into the PVC plug.

A.3.2.5 Latex Rubber Mold

A.3.2.5.1 If you plan to construct a substantial number of exposure test beakers, as described in Benoit et al. (1993), then it would be to your advantage to make a latex rubber mold to give support to the underside of the glass when drilling holes. It significantly reduces the number of broken beakers. Liquid latex, with hardener that can be purchased from the local hardware store is commonly used to coat the handles of tools. The rubber mold is constructed as follows:

1. Mix latex with hardener as per instructions.

2. Fill one exposure test beaker with the mixture.

3. Suspend one 5-cm eye bolt (5-mm diameter) with nut on end so that the eye is protruding just above the top of the mixture.

4. Allow the latex plenty of time to “set up.”

5. With proper eye protection and wearing heavy gloves, gently break the beaker with a small hammer and remove all of the glass from the mold.

6. Using a long drill bit for wood, drill an air vent hole through the mold from top through bottom.

7. When using the mold, wet the mold and the beaker with water before inserting. Place the beaker, with pre-marked location of holes, on its side in a 3.5-L stainless steel pan filled with coolant water so that the beaker is just below the surface. The beaker is then held in position with one hand while the other hand operates the drill press. Operator should wear proper eye protection.

8. After the two holes are drilled, the mold can be easily removed, with some effort, by inserting the eye bolt into the handle of a securely attached “C” clamp and physically pulling the beaker from the mold.

A.3.3 Suggested Options for More Exposure Treatments (examples given are for a three-unit treatment system)

A.3.3.1 Exposure Holding Tanks and Head Tanks

A.3.3.1.1 Same dimensions as for two-unit system except that three (3) of each should be made.

A.3.3.2 Head Tank Support Stand (1) (Figure A.3)

1. Outer diameter: 16.7 cm wide x 49.5 cm long x 17.8 cm high

2. Cutting dimensions: (acrylic plastic, 6 mm)

1 Bottom:	16.7 cm x 49.5 cm
2 Sides:	17.2 cm x 48.3 cm
2 Ends:	17.2 cm x 16.7 cm

3. Size is such that the three head tanks will fit into the support stand for storage and transport.

A.3.3.3 Water Bath (1) (Figure A.3)

1. Outer diameter: 33 cm wide x 56.4 cm long x 7.4 cm high

2. Cutting dimensions: (acrylic plastic, 6 mm)

1 Bottom:	33 cm x 71.7 cm
2 Ends:	33 cm x 6.8 cm
2 Sides:	55 cm x 6.8 cm

3. Holes: All hole sizes and locations are the same as for the two-unit system except that overflow drain is located 25.7 cm from right edge of 55-cm side. Also, two optional 1.6-cm holes centered 2.5 cm from bottom edge of 33-cm-long end piece and 1.8 cm

from corner edges are shown in the drawing for future additions of "add-on" water baths.

4. Motor mount and junction box installations are the same as for two-unit system.

A.3.3.4 "Add-on" Water Bath (example given is for one additional unit treatment system; Figure A.3)

1. Outer diameter: 18.5 cm wide x 33 cm long x 8 cm high

2. Cutting dimensions: (acrylic plastic, 6 mm)

1 Bottom:	18.5 cm x 33 cm
2 Ends:	17.3 cm x 7.4 cm
2 Sides:	33 cm x 7.4 cm

3. Holes: Inlet and outlet holes (1.6 cm) are centered 2.5 cm from bottom edge of 33-cm long side piece and 1.8 cm from corner edges.

4. The above holes will match the previously drilled holes in the main water bath. The "add-on" water bath is connected using #2 stoppers and 6.4-cm lengths of clear plastic tubing (1.3-cm diameter). The circulator pump outlet tubing (Tygon®) in the main water bath is extended through the inlet connection as shown in Figure A.2. Circulating water is then forced into the "add-on" bath and flows back to the main water bath by gravity.

5. Note that the walls of the "add-on" bath are 6 mm higher than the main water bath to accommodate the small head of water that builds up.

6. "Add-on" water baths tend to run a little warmer (0.2°C) than main water bath test temperatures.

A.3.4 Optional Automated Water-delivery Apparatus for Table Top STIR Systems (examples given are for a three-unit treatment system)

A.3.4.1 Support Stand

A.3.4.1.1 A stand to support the automated water-delivery apparatus, shown in Figure A.2, can be made from bolted slotted angle iron bolted with corner braces. A convenient size to construct is 30 cm wide x 85 cm long x 43 cm high. The head box in Figure A.2 sits on top of the stand, and the water distribution manifold as shown in Figure A.2 is placed directly under the top of the stand with two 1.3-cm conduit hangers. A small portion of each angle iron cross piece is cut away to allow the pipe to be clamped into the conduit hanger. This also keeps the manifold up high enough for sufficient clearance between the head tanks and the 6-mm pipe to hose adapters as shown in Figure A.2.

A.3.4.2 Water Renewal Supply

A.3.4.2.1 If tests will be conducted in the local water supply, then the head box water inlet shown in Figure A.2 is simply plumbed into the supply line. However, if the tests are conducted with transported water or with reconstituted water, the head box water inlet can be connected to a Nalgene® drum with flexible Tygon® tubing. With a four-volume test beaker water renewal flow rate per day, both 114-L and 208-L Nalgene® drums will hold a 5-d supply for a 3-unit treatment system and a 5-unit treatment system, respectively. If the water supply drum is located below the head box, then an open air water pump such as a March® model MDXT pump (PFC Equipment Corp., Minneapolis, MN 55440) can be used between the drum and head box.

A.3.4.3 Operation of Water-delivery Apparatus

A.3.4.3.1 The head box water inlet solenoid valve (Figure A.2) and the open air water pump (if needed) are connected to the same timer control switch. The head box water outlet solenoid valve is connected to another separate timer control switch. With four test beaker renewals/d and a 3-unit treatment system, the head box toilet float valve is pre-adjusted to allow the head box to fill to the 12-L mark on the sight tube (Figure A.2).

A.3.4.3.2 With head box filled, the renewal cycle begins when the first timer opens the head box outlet solenoid valve. The distribution manifold is quickly flooded and the 12 L of renewal water divided equally to each of the three 4-L head tanks. Since the timers have a minimum setting of one hour on-off periods, the first timer is set to shut off the head box outlet solenoid valve one hour after it opens.

A.3.4.3.3 About 30 min later, the second timer is set to open the head box water inlet solenoid valve (and pump if needed). As head box water volume reaches the 12-L mark, the pre-adjusted toilet tank valve stops the water flow. One hour after they come on, the second timer will shut off the solenoid valve inlet and water pump.

A.3.4.3.4 The automated system is then ready for the next renewal cycle that is set to begin 12 h after the first cycle. Head box volume dimensions are such that up to five-unit treatment systems can be tested simultaneously as shown in Figure A.2.

A.3.5 A criticism of the system described by Benoit et al. (1993) is that the (up to) 8 beakers placed in each holding tank are not true replicates because of the potential for exchange of water overlying the sediments among the beakers. However, this concern is largely semantic with regard to actual test results. The rationale for this position is described below. The data described below are unpublished data from USEPA Duluth (G.T. Ankley, USEPA, Duluth, MN, personal communication).

A.3.5.1 Beakers within a test tank should contain an aliquot of the *same* homogenized sediment and the same test species. The replication is intended to reflect variability

in the biology (e.g., health) of the organism, as well as placement and recovery of the animals from the test sediments (i.e., operator variability). To treat even completely separate tanks containing homogenized sediment from the same source as true replicates (of the sediment "treatment") is inaccurate and is pseudoreplication. Hence, because the same sediment is tested in each beaker in a particular tank, and because the replication is focused on defining variability in the biology of the organism (and the operator), this is essentially a nonissue from a theoretical standpoint.

A.3.5.2 From a practical standpoint, it is important to determine the potential influence of one beaker on another over the course of a test. To determine this, a study was designed (which is not advocated) in which treatments were mixed within a tank. In the first experiment, four beakers of highly metal-contaminated sediment from the Keweenaw Waterway, MI, were placed in the same tank as four beakers containing clean sediment from West Bearskin Lake, MN. This was done in two tanks; in one tank, 10 amphipods *(Hyalella azteca)* were added to each beaker, while in the other tank, 10 midges *(Chironomus tentans)* were placed in each beaker. Controls for the experiment consisted of the West Bearskin sediments assayed in separate "clean" tanks. The four contaminated beakers were placed "upstream" of the four clean beakers to attempt to maximize possible exchange of contaminant. At the end of the test, organism survival (and growth for *C. tentans)* was measured in two of the beakers from each site and sediment Cu concentrations were determined in the other two beakers from each site. The Keweenaw sediments contained concentrations of Cu in excess of 9,000 µg/g (dry wt), and were toxic to both test species (Table A.1). Conversely, survival of both *C. tentans* and *H. azteca* was high in the West Bearskin sediments from the Keweenaw tank, and was similar to survival in West Bearskin sediments held in separate tanks. Most important, there was no apparent increase in Cu concentrations in the West Bearskin sediments held in the Keweenaw tank (Table A.1).

A.3.5.3 A similar design was used to determine transfer of contaminants among beakers containing sediments spiked with the organochlorine pesticide dieldrin. In this experiment, sediment from Airport Pond, MN, was spiked with dieldrin and placed in the same tank as clean unspiked Airport Pond sediments. Two different concentrations were assayed as follows: (1) in the midge test, sediment concentrations were about 150 µg dieldrin/g (dry weight) and (2) in the amphipod test, sediments contained in excess of 450 µg dieldrin/g sediment. The control for the experiment again consisted of clean Airport Pond sediment held in a separate tank. The spiked sediments were toxic to both test species, and survival of organisms held in the clean Airport Pond sediments was similar in the two different tanks. However, there was an effect on the growth of *C. tentans* from the clean Airport Pond sediment assayed in the tank containing the spiked sediment. This corresponded to the presence of measurable dieldrin concentrations in unspiked Airport Pond sediments in the tank with the mixed treatments (Table A.2). The concentrations of dieldrin in the unspiked sediment, although detectable, were on the order of 5,000-fold lower than the spiked sediments, indicating relatively minimal transfer of pesticide.

A.3.5.4 Using a similar design, an investigation was made to evaluate if extremely low dissolved oxygen (DO) concentrations, due to sediment oxygen demand, in four beakers in a test system would result in a decrease in DO in other beakers in the tank. In this experiment, trout chow was added to each of four beakers containing clean Pequaywan Lake sediment, and placed in a test tank with four beakers containing Pequaywan Lake sediment without exogenous organic carbon. Again, the control consisted of Pequaywan Lake sediment held in a separate tank under otherwise identical test conditions. Assays were conducted, without organisms, for 10 d. At this time, DO concentrations were very low in the beakers containing trout chow-amended sediment (ca., 1 mg/L, n = 4). However, overlying water DO concentrations in the

Table A.1 Sediment Copper Concentrations and Organism Survival and Growth at the End of a 10-d Test with West Bearskin Sediment in an Individual Tank Versus 10-d Cu Concentrations and Organism Survival and Growth in West Bearskin Sediment Tested in the Same Tank as Keweenaw Waterway Sediment[1]

Sediment	Tank	Species	Survival (%)	Dry wt (mg/organism)	Cu (µg/g)
WB[2]	1	Amphipod	90	ND[3]	22.4
WB	2	Amphipod	100	ND	13.8
KW[4]	2	Amphipod	20	ND	9397.0
WB	3	Midge	95	1.34	12.3
WB	4	Midge	100	1.33	15.6
KW	4	Midge	5	ND	9167.0

[1] All values are the mean of duplicate observations (G.T. Ankley, USEPA, Duluth, MN, unpublished data)
[2] West Bearskin
[3] Not determined
[4] Keweenaw Waterway

Table A.2 Sediment Dieldrin Concentrations and Organism Survival and Growth at the End of a 10-d Test with Airport Pond Sediment in an Individual Tank Versus 10-d Dieldrin Concentrations and Organism Survival and Growth in Airport Pond Sediment Tested in the Same Tank as Dieldren-spiked Airport Pond Sediment[1]

Sediment	Tank	Species	Survival (%)	Dry wt (mg/organism)	Dieldrin (µg/g)
AP[2]	1	Amphipod	75	ND[3]	<0.01
AP	2	Amphipod	80	ND	0.07
DAP[4]	2	Amphipod	20	ND	446.4
AP	3	Midge	85	1.71	<0.01
AP	4	Midge	85	0.13	0.04
DAP	4	Midge	0	ND	151.9

[1] All values are the mean of duplicate observations (G.T. Ankley, USEPA, Duluth, MN, unpublished data)
[2] Airport Pond
[3] Not determined
[4] Dieldren-spiked Airport Pond

"untreated" vs. the "treated" beakers in a separate tank were similar, i.e., 6.8 vs. 6.9 mg/L, respectively. This indicates that from a practical standpoint, even under extreme conditions of mixed treatments (which again, is *not* recommended), interaction between beakers within a tank is minimal.

A.3.5.5 One final observation germane to this issue is worth noting. If indeed beakers of homogenized sediment within a test tank do not serve as suitable replicates, this should be manifested by a lack of variability among beakers with regard to biological assay results. This has not proven to be the case. For example, in a recent amphipod test with a homogenized sediment from the Keweenaw Waterway in which all eight replicates were held in the same tank, mean survival for the test was 76%; however, survival in the various beakers ranged from 30 to 100%, with a standard deviation of 21%. Clearly, if the test system were biased so as to reduce variability (i.e., result in unsuitable replicates due to common overlying water), this type of result would not be expected.

A.3.5.6 In summary, in both a theoretical and practical sense, use of the system described by Benoit et al. (1993) results in valid replicates that enable the evaluation of variability due to factors related to differences in organism biology and operator effects. To achieve this, it is important that treatments not be mixed within a tank; rather, the replicates should be generated from the same sediment sample. Given this, and the fact that it is difficult to document interaction between beakers using even unrealistic (and unrecommended) designs, leads to the conclusion that variability of replicates from the test system can be validly used for hypothesis testing.

A.4 Zumwalt et al. (1994) also describe a water-delivery system that can accurately deliver small volumes of water (50 mL/cycle) to eight 300-mL beakers to conduct sediment tests. The system was designed to be comparable with the system described by Benoit et al. (1993). This water-delivery system has been used in a variety of applications (i.e., Kemble et al., 1998a,b; Ingersoll et al., 1998).

A.4.1 Eight 35-mL polypropylene syringes equipped with 18-gauge needles are suspended from a splitting chamber (Figure A.4). The system is suspended above eight beakers and about 1 L of water/cycle is delivered manually or automatically to the splitting chamber. Each syringe fills and empties 50 mL into each beaker and the 600 mL of excess water empties out an overflow in the splitting chamber (Section A.4.3.1). The volume of water delivered per day can be adjusted by changing either the cycling rate or the size of the syringes. The system has been used to renew overlying water in whole-sediment toxicity tests with *H. azteca* and *C. tentans*. Variation in delivery of water among 24 beakers was less than 5%. The system is inexpensive (<$100), easy to build (<8 h), and easy to calibrate (<15 min).

A.4.2 Water-Splitting Chamber

A.4.2.1 The glass water-splitting chamber is 14.5 cm wide, 30 cm long, and 6.5 cm high (inner diameter). Eight 3.8-cm holes and one 2.5-cm hole are drilled in a 15.5 cm x 30.5 cm glass bottom before assembly (Figure A.4 and Table A.3). The glass bottom is made from 4.8- (3/16 inch) or 6.4-mm (1/4 inch) plate glass. An easy way to position the 3.8-cm holes is to place the eight 300-mL beakers (2 wide x 4 long) under the bottom plate and mark the center of each beaker. The 2.5-cm hole for overflow is centered at one end of the bottom plate between the last two holes and endplate (Figure A.4). After drilling the holes in the bottom plate, the side (6.5 x 30.5 cm) and end (6.5 x 14.5 cm) plates are cut from 3.2-mm (1/8 inch) double-strength glass and the splitting box is assembled using silicone adhesive. Sharp glass edges should be sanded smooth using a whetstone or a piece of carborundum wheel. After the splitting chamber has dried for 24 h, four 12-mm (outer diameter) stainless-steel tubes (7 cm long) are glued to each corner of the splitting chamber (the surface of the steel tubes is scored with rough emery paper to allow better adhesion of the silicone). These tubes are used as sleeves for attaching the legs to the splitting chamber. The legs of the splitting chamber are threaded stainless-steel rods (9.5 mm [3/8 inch] diameter, 36 cm long). The location of the tubes depends on the way that the beakers are to be accessed in the waterbath. If the tubes are placed on the side of the splitting chamber, a 3.2-mm-thick x 2-cm-wide x 7-cm-long spacer is required so beakers and the optional waterbath can be slid out the ends (Figure A.4). If the sleeves and legs are attached to the ends of the splitting chamber, the beakers and waterbath can be removed from the side. The legs are inserted into the 12-mm tubes and secured using nylon nuts or wingnuts. The distance between the tips of the needles to the surface of the water in the 300-mL beakers is about 2 cm. Four 1-L beakers could also be placed under the splitting chamber.

A.4.2.2 A #7 silicone stopper drilled with a 21-mm (outer diameter) core borer is used to hold each 35-mL polypropylene syringe (45 mL total capacity) in place. Glass syringes could be used if adsorption of contaminants on the surface of the syringe is of concern. A dilute soap solution can be used to help slide the syringe into the #7 stopper (until the end of the syringe is flush with the top of stopper). Stoppers and syringes are inserted into 3.8-cm holes and are visually leveled. A #5 silicone stopper drilled with an 8-mm (outer diameter) core borer is placed in the 2.5 cm overflow hole. An 8-mm (outer diameter) glass tube (7.5 cm long) is inserted into the stopper. Only 3 mm of the overflow tube should be left exposed above the stopper. This overflow drain is placed about 3 mm lower than the top of the syringes. A short piece of 6.4-mm (1/4 inch; inner diameter) tubing can be placed on the lower end of drain to collect excess water from the overflow.

A.4.2.3 The splitting chamber is leveled by placing a level on top of the chamber and adjusting the nylon nuts. Eighteen-gauge needles are attached to the syringes.

Figure A.4 Water splitting chamber described in Zumwalt et al. (1994).

Table A.3 Materials Needed for Constructing a Zumwalt et al. (1994) Delivery System

Equipment
Drill press
Glass drill bits (2.54 cm [1 inch] and 3.8 cm [1.5 inch])
Cork boring set
Table-top saw equipped with a carborundum wheel
Small level (about 30 cm long)
Supplies
300-mL beakers (lipless, tall form; e.g., Pyrex Model 1040)
Stainless-steel screen (50- x 50-mesh)
9.5-mm (3/8 inch x 16) stainless-steel threaded rod
9.5-mm (3/8 inch x 16) nylon wingnuts
9.5-mm (3/8 inch x 16) nylon nuts
35-mL Mono-ject syringes (Sherwood Medical, St. Louis, MO)
18-gauge Mono-ject stainless-steel hypodermic needles
Silicone stoppers (#0, 5, and 7)
Plate glass (6.4 mm [1/4 inch], 4.8 mm [3/16 inch], 3.2 mm [1/8 inch])
Glass tubing (8-mm outer diameter)
Stainless-steel tubing (12-mm outer diameter)
Silicone adhesive (without fungicide)
5-way stainless-steel gang valves and Pasteur pipets (14.5 cm [5.75 inch])

About 6 mm of the needle should remain after the sharp tip has been cut off using a carborundum wheel. Jagged edges left in the bore of the needle can be smoothed using a small sewing needle or stainless-steel wire.

A.4.2.4 When about 1 L of water is delivered to the splitting chamber, the top of each syringe should be quickly covered with water. The overflow tube will quickly drain excess water to a level just below the tops of the syringes. The syringes should empty completely in about 4 min. If water remains in a syringe, the needle should be checked to ensure that it is clean and does not have any jagged edges.

A.4.3 Calibration and Delivery of Water to the Splitting Chamber

A.4.3.1 Flow adjustments can be made by sliding either the stoppers or syringes up or down to deliver more or less water. A splitting chamber with eight syringes can be calibrated in less than 15 min. Delivery of water to the splitting chamber can be as simple as manually adding about 1 L of water/cycle. Water can be added automatically to the splitting chamber using a single cell or a Mount and Brungs (1967) diluter that delivers about 1 L/cycle on a time delay. About 50 mL will be delivered to each of the 8 beakers/cycle and 600 mL will flow out the overflow. A minimum of about 1 L/cycle should be dumped into the splitting chamber to ensure each syringe fills to the top. If the quantity of water is limited at a laboratory, the excess water that drains through the overflow can be collected and recycled.

A.4.4 Waterbath and Exposure Beakers

A.4.4.1 The optional waterbath surrounding the beakers is made from 3.2-mm (1/8-inch) double-strength glass and is 15.8 cm wide x 29.5 cm long x 11.7 cm high (Figure A.4 [Figure A.3 in the Benoit et al., 1993 system]). Before the pieces are assembled, a 1.4-cm hole is drilled in one of the end pieces. The hole is 7.2 cm from the bottom and centered between each side of the end piece. A glass tube inserted through a #0 silicone stopper can be used to drain water from the waterbath. A notch is made in each 300-mL beaker by making two cuts with a carborundum wheel 1.9 cm apart to the 275 mL level. The beaker is etched across the bottom of the cuts, gently tapped to remove the cut section, and the notch is covered with 50- x 50-mesh stainless-steel screen using silicone adhesive. The waterbath illustrated in Figure A.4 is optional if the splitting chambers and beakers are placed in a larger waterbath to collect waste water. This smaller waterbath could be used to collect waste water and a surrounding larger waterbath could be used for temperature control.

A.4.5 Operation and Maintenance

A.4.5.1 Maintenance of the system is minimal. The syringes should be checked daily to make sure that all of the water is emptying with each cycle. As long as the syringe empties completely, the rate of flow out of the syringes is not important because a set volume of water is delivered from each syringe. If the syringe does not empty completely with each cycle, the needle tip should be replaced or cleaned with a thin wire or sewing needle. If the screens on the beakers need to be cleaned, a toothbrush can be used to brush the outside of screens.

A.4.5.2 Overlying water can be aerated by suspending Pasteur pipets (e.g., Pyrex disposable 14.5-cm [5.75 inch] length) about 3 cm above the sediment surface in the beakers. Five-way stainless-steel gang valves are suspended from the splitting chamber using stainless-steel hooks. Latex tubing (3.2-mm [1/8 inch] inner diameter) is used to connect valves and pipets. Flow rate of air should be maintained at about 2 to 3 bubbles/s and the pipets can be placed on the outside of the beakers when samples of overlying water are taken during a test.

A.4.5.3 The splitting chambers were used to deliver water in a toxicity test with the midge *Chironomus tentans* exposed to metal-contaminated sediments (Zumwalt et al., 1994). Ten third-instar midges were exposed in 300-mL beakers containing 100 mL of sediment and 175 mL of overlying water at 23°C. Midges in each beaker received a daily suspension of 4 mg Tetrafin® flake food and survival and growth were measured after 10 d. Splitting chambers delivered 50 mL/cycle of overlying water to each of the eight replicate beakers/sediment sample. One liter of water was delivered with a single-cell diluter to each splitting chamber 4 times/d. This cycle rate resulted in 1.1 volume additions of overlying water/d to each beaker ([4 cycles/d x 50-mL volume/cycle]/175 mL of overlying water). The variation in delivery of water between 24 beakers was less than 5%.

A.4.5.4 Hardness, alkalinity, and conductivity in water overlying the sediments averaged about 20% higher than inflowing water. These water-quality characteristics tended to be more similar to inflowing water at the end of the

exposure compared with the beginning of the exposure. The average pH was about 0.3 units lower than inflowing water. Ammonia in overlying water ranged from 0.20 to 0.83 mg/L. The dissolved oxygen content was about 1 mg/L lower than inflowing water at the beginning of the exposure and was about 2 to 3 mg/L lower than inflowing water by the end of the exposure. Survival and growth of midges were reduced with exposure to metal-contaminated sediments. Water delivered at a similar rate to a second set of beakers using a system described by Benoit et al. (1993) resulted in similar overlying water quality and similar toxic effects on midges.

A.4.5.5 The system has been used to deliver 33 ‰ salt water to exposure chambers for 10 d. Precipitation of salts on the tips of the needles reduced flow from the syringes. Use of a larger bore needle (16-gauge) reduced clogging problems; however, daily brushing of the needle tips is required. Use of larger bore needles with 300-mL beakers containing 100 mL of sediment and 175 mL of overlying water results in some suspension of sediment in the overlying water. This suspension of sediment can be eliminated if the stream of water from the larger bore needle falls on a baffle (e.g., a piece of glass) at the surface of the water in the beaker.

A.5 Brunson et al. (1998) describe a water-delivery system for use with larger exposure chambers in the *Lumbriculus variegatus* sediment exposures (Section 13). Exposures of oligochaetes by Brunson et al. (1998) were conducted for 28 d in 4-L glass beakers containing 1 L of sediment and 3 L of overlying water. Four replicate chambers were tested for each sediment sample evaluated. Each beaker was calibrated to 4 L using a glass standpipe that exited through the beaker wall and was held in place with a silicon stopper. Beakers received 2 volume additions (6 L) of overlying water per day. Water was delivered using a modified Mount and Brungs diluter system that was designed to deliver 1 L/cycle (Ingersoll and Nelson, 1990). An in-line flow splitter was attached to each delivery line to split the water flow evenly to each of four beakers. These splitters were constructed of 1/4 inch PVC pipe with four silicone stoppers and 14-gauge stainless-steel hypodermic needles with the points and connector ends cut off the needles (Figure A.5). Glass stands were used to support the splitters, keeping them level to maintain a constant volume delivery to each beaker (+ 5%).

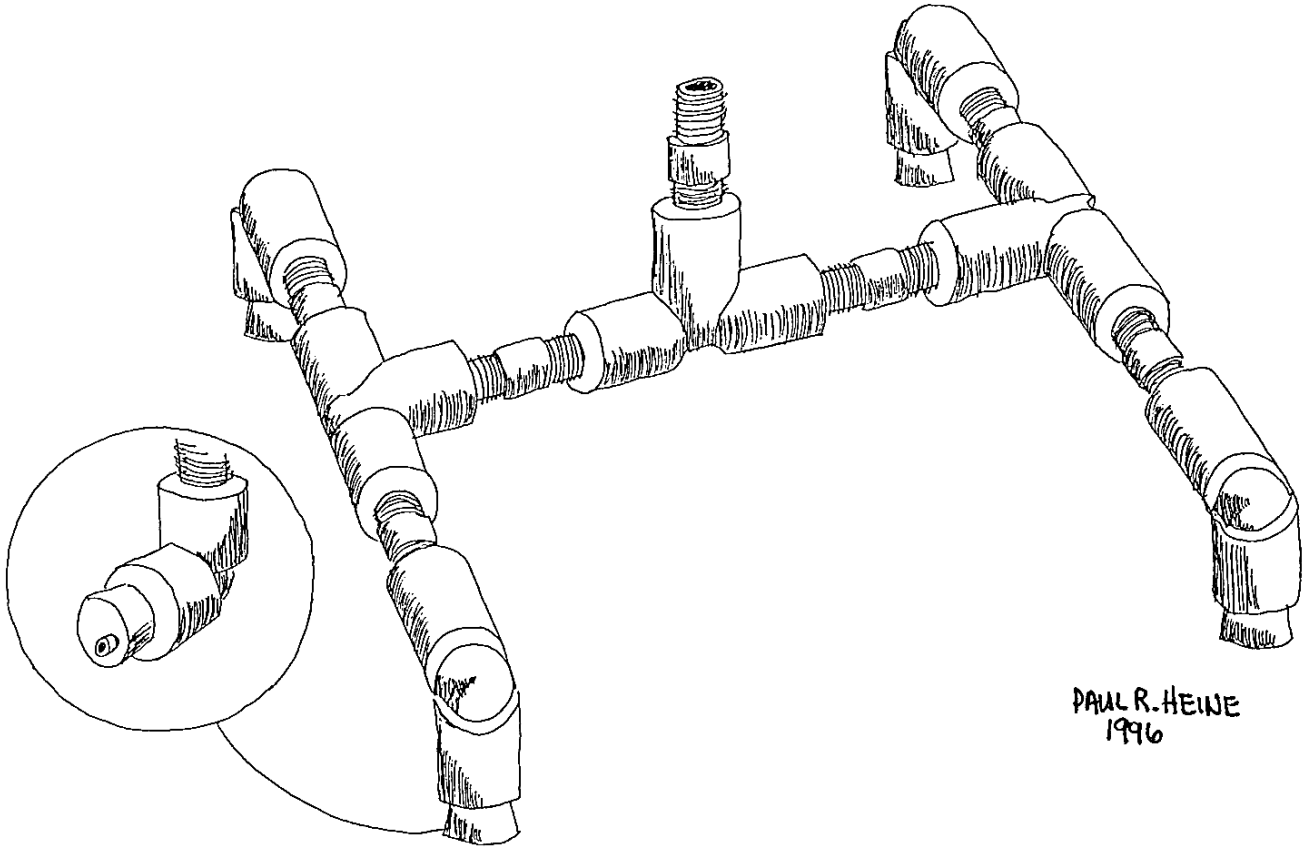


Figure A.5. Diagram of in-line flow splitter used to deliver overlying water in the sediment exposures of *Lumbriculus variegatus* (Brunson et al., 1998).

Appendix B
Food Preparation

B.1 Yeast, Cerophyl®, and Trout Chow (YCT) for Feeding the Cultures and *Hyalella azteca*

B.1.1 Food should be stored at 4°C and used within two weeks from preparation; however, once prepared, YCT can be frozen until use.

B.1.2 Digested trout chow is prepared as follows:

1. Preparation of trout chow requires one week. Use 1/8 inch pellets prepared according to current U.S. Fish and Wildlife Service specifications. Suppliers of trout chow include Zeigler Bros., Inc., P.O. Box 95, Gardners, PA, 17324 (717/780-9009); Glencoe Mills, 1011 Elliott, Glencoe, MN, 55336 (320/864-3181); and Murray Elevators, 118 West 4800 South, Murray, UT 84107 (800/521-9092).

2. Add 5.0 g of trout chow pellets to 1 L of deionized water. Mix well in a blender and pour into a 2-L separatory funnel or similar container. Digest before use by aerating continuously from the bottom of the vessel for one week at ambient laboratory temperature. Water lost due to evaporation is replaced during digestion. Because of the offensive odor usually produced during digestion, the vessel should be placed in a ventilated area.

3. At the end of the digestion period, allow material to settle for a minimum of 1 h. Filter the supernatant through a fine mesh screen (e.g., Nitex® 110 mesh). Combine with equal volumes of the supernatant from Cerophyl® and yeast preparation (below). The supernatant can be used fresh, or it can be frozen until use. Discard the remaining particulate material.

B.1.3 Yeast is prepared as follows:

1. Add 5.0 g of dry yeast, such as Fleishmann's® Yeast, Lake State Kosher Certified Yeast, or equivalent, to 1 L of deionized water.

2. Stir with a magnetic stirrer, shake vigorously by hand, or mix with a blender at low speed, until the yeast is well dispersed.

3. Combine the yeast suspension immediately (do not allow to settle) with equal volumes of supernatant from the trout chow (above) and Cerophyl® preparations (below). Discard excess material.

B.1.4 Cerophyl® is prepared as follows:

1. Place 5.0 g of dried, powdered cereal or alfalfa leaves, or rabbit pellets, in a blender. Cereal leaves are available as "Cereal Leaves" from Sigma Chemical Company, P.O. Box 14508, St. Louis, MO, 63178 (800/325-3010); or as Cerophyl®, from Ward's Natural Science Establishment, Inc., P.O. Box 92912, Rochester, NY, 14692-9012 (716/359-2502). Dried, powdered alfalfa leaves may be obtained from health food stores, and rabbit pellets are available at pet shops.

2. Add 1 L of deionized water.

3. Mix in a blender at high speed for 5 min, or stir overnight at medium speed on a magnetic stir plate.

4. If a blender is used to suspend the material, place in a refrigerator overnight to settle. If a magnetic stirrer is used, allow to settle for 1 h. Decant the supernatant and combine with equal volumes of supernatant from trout chow and yeast preparations (above). Discard excess material.

B.1.5 Combined yeast-Cerophyl-trout chow (YCT) is mixed as follows:

1. Thoroughly mix equal (e.g., 300 mL) volumes of the three foods as described above.

2. Place aliquots of the mixture in small (50 mL to 100 mL) screw-cap plastic bottles.

3. Freshly prepared food can be used immediately, or it can be frozen until needed. Thawed food is stored in the refrigerator between feedings and is used for a maximum of two weeks. Do not store YCT frozen over three months.

4. It is advisable to measure the dry weight of solids in each batch of YCT before use. The food should contain 1.7 to 1.9 g solids/L.

B.2 Algal Food

B.2.1 Starter cultures of the green algae, *Selenastrum capricornutum* are available from the following sources: American Type Culture Collection (Culture No. ATCC 22662), 12301 Parklawn Drive, Rockville, MD 10852, or Culture Collection of Algae, Botany Department, University of Texas, Austin, TX 78712.

B.2.2 Algal Culture Medium for the green algae is prepared as follows (USEPA, 1993a):

1. Prepare stock nutrient solutions using reagent grade chemicals as described in Table B.1.

Table B.1 Nutrient Stock Solutions for Maintaining Algal Stock Cultures

Stock solution		Compound	Amount dissolved in 500 mL deionized water
1.	Macronutrients		
	A.	$MgCl_2 \cdot 6H_2O$	6.08 g
		$CaCl_2 \cdot 2H_2O$	2.20 g
		$NaNO_3$	12.75 g
	B.	$MgSO_4 \cdot 7H_2O$	7.35 g
	C.	K_2HPO_4	0.522 g
	D.	$NaHCO_3$	7.50 g
2.	Micronutrients		
		H_3BO_3	92.8 mg
		$MnCl_2 \cdot 4H_2O$	208.0 mg
		$ZnCl_2$	l.64 mg[1]
		$FeCl_3 \cdot 6H_2O$	79.9 mg
		$CoCl_2 \cdot 6H_2O$	0.714 mg[2]
		$Na_2MoO_4 \cdot 2H_2O$	3.63 mg[3]
		$CuCl_2 \cdot 2H_2O$	0.006 mg[4]
		$Na_2EDTA \cdot 2H_2O$	l50.0 mg
		Na_2SeO_4	1.196 mg[5]

[1]$ZnCl_2$—Weigh out 164 mg and dilute to 100 mL. Add 1 mL of this solution to micronutrient stock.

[2]$CoCl_2 \cdot 6H_2O$—Weigh out 71.4 mg and dilute to 100 mL. Add 1 mL of this solution to micronutrient stock.

[3]$Na_2MoO_4 \cdot 2H_2O$—Weigh out 36.6 mg and dilute to 10 mL. Add 1 mL of this solution to micronutrient stock.

[4]$CuCl_2 \cdot 2H_2O$—Weigh out 60.0 mg and dilute to 1000 mL. Take 1 mL of this solution and dilute to 10 mL. Take 1 mL of the second dilution and add to micronutrient stock.

[5]Na_2SeO_4—Weigh out 119.6 mg and dilute to 100 mL. Add 1 mL of this solution to micronutrient stock.

2. Add 1 mL of each stock solution, in the order listed in Table B.1, to about 900 mL of deionized water. Mix well after the addition of each solution. Dilute to 1 L, mix well. The final concentration of macronutrients and micronutrients in the culture medium is listed in Table B.2.

3. Immediately filter the medium through a 0.45 µm pore diameter membrane at a vacuum of not more than 380 mm (15 in.) mercury, or at a pressure of not more than one-half atmosphere (8 psi). Wash the filter with 500 mL deionized water before use.

4. If the filtration is carried out with sterile apparatus, the filtered medium can be used immediately, and no further sterilization steps are required before the inoculation of the medium. The medium can also be sterilized by autoclaving after it is placed in the culture vessels. Unused sterile medium should not be stored more than one week before use, because there may be substantial loss of water by evaporation.

B.2.3 Algal Cultures

B.2.3.1 Two types of algal cultures are maintained: (1) stock cultures and (2) “food” cultures.

Table B.2 Final Concentration of Macronutrients and Micronutrients in the Algal Culture Medium

Macronutrient	Concentration (mg/L)	Element	Concentration (mg/L)
$NaNO_3$	25.5	N	4.20
$MgCl_2 \cdot 6H_2O$	12.2	Mg	2.90
$CaCl_2 \cdot 2H_2O$	4.41	Ca	1.20
$MgSO_4 \cdot 7H_2O$	14.7	S	1.91
K_2HPO_4	1.04	P	0.186
$NaHCO_3$	15.0	Na	11.0
		K	0.469
		C	2.14

Micronutrient	Concentration (µg/L)	Element	Concentration (µg/L)
H_3BO_3	185	B	32.5
$MnCl_2 \cdot 4H_2O$	416	Mn	115
$ZnCl_2$	3.27	Zn	1.57
$CoCl_2 \cdot 6H_2O$	1.43	Co	0.354
$CuCl_2 \cdot 2H_2O$	0.012	Cu	0.004
$Na_2MoO_4 \cdot 2H_2O$	7.26	Mo	2.88
$FeCl_3 \cdot 6H_2O$	160	Fe	33.1
$Na_2EDTA \cdot 2H_2O$	300	—	—
Na_2SeO_4	2.39	Se	0.91

B.2.3.2 Establishing and Maintaining Stock Cultures of Algae

1. Upon receipt of the "starter" culture of *S. capricornutum* (usually about 10 mL), a stock culture is started by aseptically transferring 1 mL to each of several 250-mL culture flasks containing 100 mL algal culture medium (prepared as described above). The remainder of the starter culture can be held in reserve for up to six months in a refrigerator (in the dark) at 4°C.

2. The stock cultures are used as a source of algae to initiate "food" cultures. The volume of stock culture maintained at any one time will depend on the amount of algal food required for culture. Stock culture volume may be rapidly "scaled up" to several liters using 4-L serum bottles or similar vessels containing 3 L of growth medium.

3. Culture temperature is not critical. Stock cultures may be maintained at 25°C in environmental chambers with cultures of other organisms if the illumination is adequate (continuous "cool-white" fluorescent lighting of about 4300 lux).

4. Cultures are mixed twice daily by hand.

5. Stock cultures can be held in the refrigerator until used to start "food" cultures, or can be transferred to new medium weekly. One to 3 mL of 7-d-old algal stock culture, containing about 1.5 X 10^6 cells/mL are transferred to each 100 mL of fresh culture medium. The inoculum should provide an initial cell density of about 10,000 to 30,000 cells/mL in the new stock cultures. Aseptic techniques should be used in maintaining the stock algal cultures, and care should be exercised to avoid contamination by other microorganisms.

6. Stock cultures should be examined microscopically weekly at transfer for microbial contamination. Reserve quantities of culture organisms can be maintained for 6 to 12 months if stored in the dark at 4°C. It is advisable to prepare new stock cultures from "starter" cultures obtained from established outside sources of organisms every four to six months.

B.2.3.3 Establishing and Maintaining "S. capricornutum Food" Cultures

1. "*S. capricornutum* food" cultures are started 7 d before use. About 20 mL of 7-d-old algal stock culture (described in the previous paragraph), containing 1.5 X 10^6 cells/mL are added to each liter of fresh algal culture medium (e.g., 3 L of medium in a 4-L bottle or 18 L in a 20-L bottle). The inoculum should provide an initial cell density of about 30,000 cells/mL. Aseptic techniques should be used in preparing and maintaining the cultures, and care should be exercised to avoid contamination by other microorganisms. However, sterility of food cultures is not as critical as in stock cultures because the food cultures are used in 7 to 10 d. A one-month supply of algal food can be grown at one time and stored in the refrigerator.

2. Food cultures may be maintained at 25°C in environmental chambers with the algal stock cultures or cultures of other organisms if the illumination is adequate (continuous "cool-white" fluorescent lighting of about 4300 lux).

3. Cultures are mixed continuously on a magnetic stir plate (with a medium size stir bar), in a moderately aerated separatory funnel, or are manually mixed twice daily. If the cultures are placed on a magnetic stir plate, heat generated by the stirrer might elevate the culture temperature several degrees. Caution should be taken to prevent the culture temperature from rising more than 2 to 3°C.

B.2.3.4 Preparing Algal Concentrate of S. capricornutum for Use as Food

1. An algal concentrate of *S. capricornutum* containing 3.0 to 3.5 X 10^7 cells/mL is prepared from food cultures by centrifuging the algae with a plankton or bucket-type centrifuge, or by allowing the cultures to settle in a refrigerator for at least one week and siphoning off the supernatant.

2. The cell density (cells/mL) in the concentrate is measured with an electronic particle counter, microscope and hemocytometer, fluorometer, or spectrophotometer and used to determine the dilution (or further concentration) required to achieve a final cell count of 3.0 to 3.5 X 10^7 cells/mL.

3. Assuming a cell density of about 1.5 X 10^6 cells/mL in the algal food cultures at 7 d, and 100% recovery in the concentration process, a 3-L culture at 7 to 10 d will provide 4.5 X 10^9 algal cells.

4. Algal concentrate can be stored in the refrigerator for one month.

5. Cultures of *Hyalella azteca* are fed 10 mL/L on renewal/harvest days and 5 mL/L on all other days (USEPA, 1993c).

B.2.3.5 Cell Counts

1. Several types of automatic electronic and optical particle counters are available to rapidly count cell number (cells/mL) and mean cell volume (MCV; μm^3/cell). The Coulter Counter is widely used and is discussed in detail in USEPA (1978). When the Coulter Counter is used, an aliquot

(usually 1 mL) of the test culture is diluted 10X to 20X with a 1% sodium chloride electrolyte solution, such as Coulter ISOTON®, to facilitate counting. The resulting dilution is counted using an aperture tube with a 100-µm diameter aperture. Each cell (particle) passing through the aperture causes a voltage drop proportional to its volume. Depending on the model, the instrument stores the information on the number of particles and the volume of each, and calculates the mean cell volume. The following procedure is used:

A. Mix the algal culture in the flask thoroughly by swirling the contents of the flask about six times in a clockwise direction, and then six times in the reverse direction; repeat the two-step process at least once.

B. At the end of the mixing process, stop the motion of the liquid in the flask with a strong brief reverse mixing action, and quickly remove 1 mL of cell culture from the flask with a sterile pipet.

C. Place the aliquot in a counting beaker, and add 9 mL (or 19 mL) of electrolyte solution (such as Coulter ISOTON®).

D. Determine the cell density (and MCV, if desired).

2. Manual microscope counting methods for cell counts are determined using a Sedgwick-Rafter, Palmer-Maloney, hemocytometer, inverted microscope, or similar methods. For details on microscope counting methods, see APHA (1992) and USEPA (1973). Whenever feasible, 400 cells per replicate are counted to obtain ±10% precision at the 95% confidence level. This method has the advantage of allowing for the direct examination of the condition of the cells.

B.3 Tetrafin® Food (or Other Fish Flake Food) for Culturing and Testing *Chironomus tentans*

B.3.1 Food should be stored at 4°C and used within two weeks from preparation or can be frozen until use. If it is frozen, it should be reblended, once thawed, to break up any clumps

1. Blend the Tetrafin® food in deionized water for 1 to 3 min or until very finely ground.

2. Filter slurry through an #110 Nitex screen to remove large particles. Place aliquot of food in 100- to 500-mL screw-top plastic bottles. It is desirable to determine dry weight of solids in each batch of food before use. Food should be held for no longer than two weeks at 4°C. Food can be frozen before use, but it is desirable to use fresh food.

3. Tetrafin® food is added to each culture chamber to provide about 0.04 mg dry solids/mL of culture water. A stock suspension of the solids is prepared in culture water such that a total volume of 5.0 mL of food suspension is added daily to each culture chamber. For example, if a culture chamber volume is 8 L, 300 mg of food would be added daily by adding 5 mL of a 56 g/L stock suspension (USEPA, 1993).

4. In a sediment test, Tetrafin® food (4.0 g/L) is added at 1.5 mL daily to each test chamber.

Appendix C
Supplies and Equipment for Conducting the *Chironomus tentans* Life-cycle Sediment Toxicity Test

C.1 General

C.1.1 Section 15 outlines the methods for conducting a *Chironomus tentans* life-cycle sediment toxicity test. This Appendix describes the equipment needed to conduct this test.

C.2 Emergence Traps (Figure C.1)

C.2.1 These traps are needed from Day 20 to the end of the test. These traps fit on the top of the lipless glass beakers with the narrow end up. These are 5-ounce plastic cups with 14-mesh nylon screen glued to the cup in place of the plastic bottom.

C.3 Reproduction/Oviposit Chambers (R/O; Figure C.2)

C.3.1 These R/O chambers use emergence traps and are needed once adults begin to emerge. Emergence traps are used to store adults collected daily, and are placed in a 100- X 20-mm petri dish that contains about 50 mL of overlying water. When emergence occurs, the emergence traps containing adults are removed and placed onto a petri dish. At least one male for each emergent female is added, and the R/O chamber (Figure C.2) is placed back into the test system or into environmental chambers maintained at the appropriate temperature and lighting. A new emergence trap is then placed on top of the lipless beaker. The R/O chambers are kept in this manner to collect the egg masses and track mortality of adults. If space is not a limiting factor, maintaining one R/O chamber per pair of organisms is encouraged. Where space is limited, many adults may be kept in a single R/O chamber, and the chambers may be double stacked (Double Stack Support Stand described in Section C.8) using a larger plastic (9-ounce) cup that serves as a stand for the second level of the emergence trap. The egg masses are removed by lifting the edge of the cup enough to permit transfer with a pipet.

C.4 Adult Collector Dish (Figure C.3)

C.4.1 This is used as a tray which is placed under the emergence trap or reproduction/oviposit (R/O) chambers to provide access to adults and to facilitate transfer of the males and females as needed. This dish is constructed of large petri dishes, i.e., 100- X 20-mm glass dishes or 100- X 20-mm plastic dishes. A 2.54-cm hole is cut in the middle and covered with 58-mesh opening nylon screen. Two slits are cut within the screen at 90 degree angles to each other. This facilitates insertion of the aspirator tube without risk of the adults flying away.

C.5 Aspirator (Figure C.3)

C.5.1 This is used to collect and transfer adults from the reproduction/oviposit (R/O) chambers. A 60-cc syringe is modified by cutting the end with the tip off and adding a retainer to hold the emergence traps and reproductive/oviposit chambers. The retainer is a 7-cm diameter plastic lid (from 270-mL wide mouth glass jar) and a large stopper is used to hold the syringe. The stopper and the lid is drilled with a hole saw of about 1 inch. The large stopper is glued to the lid. This retainer is then attached to the syringe. To facilitate transferring the animals, prepare two tubes, one about 16 cm in length and one about 4 cm (6-mm ID) and place these in a stopper (i.e., No. 5, 5.5 or 6) that has been drilled with two holes. Fasten a section (about 70 cm) of tygon tubing onto the short piece of glass and cover the tube with a piece of thin stainless steel screen (250-µm mesh) before inserting the tube into the rubber stopper. Adults should be stationary in trap to minimize the possibility of escape.

C.6 Auxiliary Male Holding Dish

C.6.1 When emergence begins in the auxiliary beakers, the males are transferred individually to inverted 60- X 15-mm plastic petri dishes with several small holes (3 mm in diameter) drilled in the top. A thin layer of overlying water (about 5 mL) is added and renewed until the males are needed for the reproduction chambers. These males are held in the test system for temperature control, and can be used for up to 5 d after collection.

C.7 Egg Hatching Chamber

C.7.1 Petri dishes, 60- X 15-mm plastic, are used to incubate (23°C) egg masses in approximately 15 mL of water. Hatch is monitored for 6 d. Hatch success is determined by subtracting the number of unhatched eggs at the end of 6 d from the initial estimate of the egg mass.

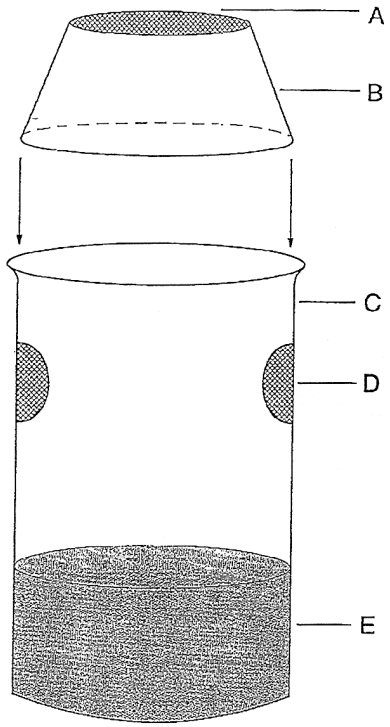


Figure C.1. Emergence trap used in the life-cycle *Chironomus tentans* sediment test. A: the nylon screen; B: the inverted plastic cups; C: the 300-mL lipless exposure beaker; D: the water exchange screen ports; E: test sediment.

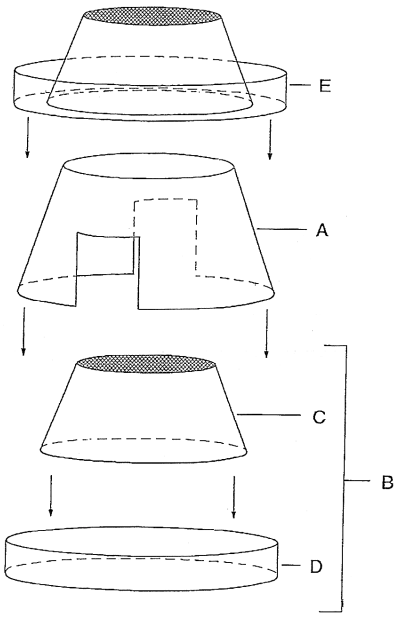


Figure C.2. The reproduction/oviposit chamber with the double stack support stand. A: the notched, inverted 270-ml (9-oz) plastic cup used to allow double stacking; B: the reproduction/oviposit (R/O) unit (C and D); C: inverted, 120-mL (4-oz) plastic cup with nylon screen; D: one-half of petri dish (100 X 20 mm) with 50 mL of overlying water; E: the reproduction/oviposit (R/O) chamber.

C.8 Supplies and Sources

A. Emergence Trap/Reproduction Oviposit Chamber.

1. 120-mL (5-ounce) plastic cups, Plastics Inc., St. Paul, MN 55164.

2. 1400-mesh opening (micron) nylon screen (mesh count = 14/inch), Monodur® 1400 Farbric Corporation, 7160 Northland Circle, Minneapolis, MN 55428, 612/535-3220.

B. Double Stack Support Stand: 270-mL (9-ounce) plastic cups, Solo Inc, Urbana, IL, 61801-2895.

C. Aspirator.

1. 60-cc syringe, 1 each, B-D® No. 309663, Becton and Dickinson & Company, Franklin Lakes, NJ 07417-1884.

2. 7-cm diameter plastic lid, 1 each.

3. Rubber stopper, 1 each, size 10, 10.5, or 11.

4. Rubber stopper, 1 each, size 5.5 or 6.

5. Glass tubing, 6-mm I.D., 1- 16 cm long, 1- 4 cm long.

6. Nalgene 6-mm plastic connector for mouth piece.

7. Stainless-steel screen, 250-µm mesh.

D. Auxillary Male Holding Chamber: 60- X 15-mm petri dish with 3-mm holes drilled, Falcon 1007 B-D®, Becton and Dickinson and Company, Franklin Lakes, NJ 07417-1884.

E. Egg Hatching Chambers: 60- X 15-mm petri dish, Falcon 1007 B-D®, Becton and Dickinson and Company, Franklin Lakes, NJ 07417-1884.

F. Adult Collector Dish:

1. 100- X 20-mm glass petri dish with a 2.54-cm access hole, Corning Glassware Corning, New York or 100- X 20-mm plastic petri dish with a 2.54-cm access hole, Falcon 1005 B-D®, Becton and Dickinson and Company, Franklin Lakes, NJ 07417-1884.

2. 58-mesh opening nylon screen, cut with slits at 90° angles to each other, Monodur®,Farbric Corporation, 7160 Northland Circle, Minneapolis, MN 55428, 612/535-3220.

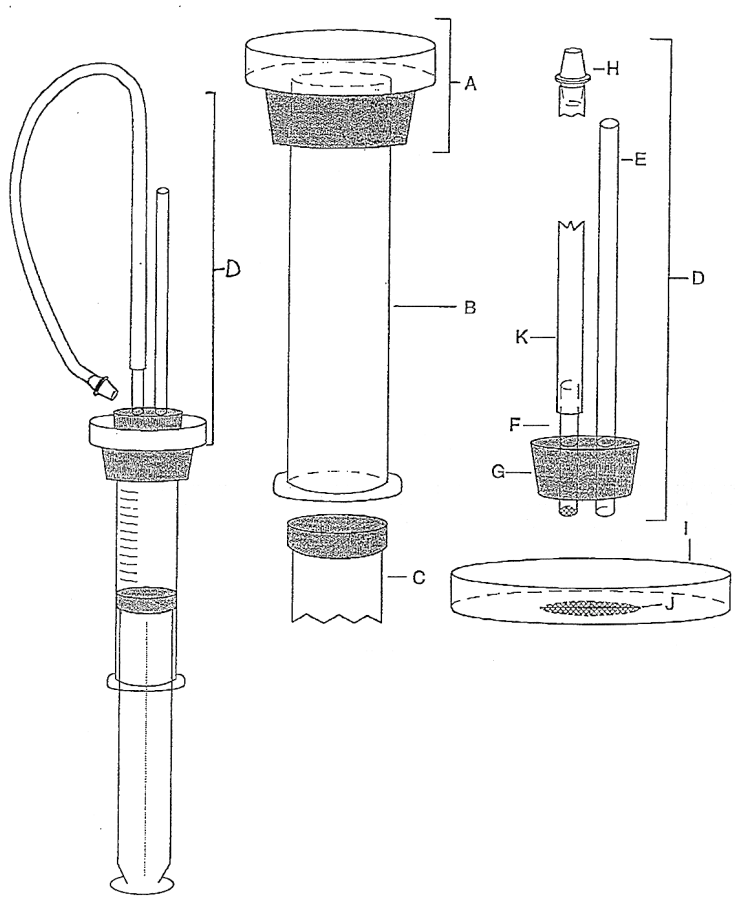


Figure C.3. Adult collection/transfer equipment. A: transfer retainer unit showing inverted plastic cover and rubber stopper glued inside of it; B: 60-cc syringe; C: plunger; D: detachable aspirator unit; E: long glass collector tube; F: short glass tube to serve as connector for inhaler tube; note stainless steel screen attached to end through stopper; G: 2-hole rubber stopper; H: nalgene plastic connector attached to tygon tubing and used as a mouthpiece to provide slight suction; I: collector dish, one-half of glass or plastic petri dish; J: petri dish with hole access that is screen covered and slotted; K: tygon tubing attached to glass tubing (F).

C.9 Construction of an Adult Midge Emergence Trap for Use in a "Zumwalt" Exposure System in Life-cycle Sediment Tests

C.9.1 The construction of the emergence trap described in Figure C.4 is an alternate design to the trap illustrated in Figures C.1 and C.2. The emergence trap illustrated in Figure C.4 is designed to fit under the exposure system described by Zumwalt et al. (1994; Section A.4). The level of the syringes will need to be raised about 1 1/2 inches using the threaded steel rods supporting the upper chamber.

C.9.2 Cut a 2 1/2-inch plexiglass tube into 1 1/4-inch-long pieces using a bandsaw or miter box and a handsaw.

C.9.3 Drill a 1/2-inch hole in the side (middle) of the 1 1/4-inch ring of plexiglass. Cut a small board to fit inside of the 1 1/4-inch ring to help support the plexiglass when drilling. The 1/2-inch drill bit should be dulled to help prevent the bit from digging in too fast.

C.9.4 Drill three 1/16-inch holes in the plexiglass ring spaced evenly around the ring and 1/4 inch off the bottom of the ring.

C.9.5 Trace around the stainless-steel screen. Cut out screen and place on top of the plexiglass ring. Use a propane-soldering torch or glass-blowing torch to heat up one end of a 1/4-inch or 3/8-inch threaded steel rod (about 12 to 15 inches long so that one end remains cool). Press the hot end of the steel rod against the screen and plexiglass until the screen melts into the plexiglass (usually a few seconds). Repeat the process until the screen is completely melted to the top of the plexiglass ring.

C.9.6. Bend 4-mm glass tubing (outer diameter) over a propane-soldering torch or glass-blowing torch and cut the tubing with a glass wheel or etch the tubing with a file to break. This glass tube is only to be used if beakers need

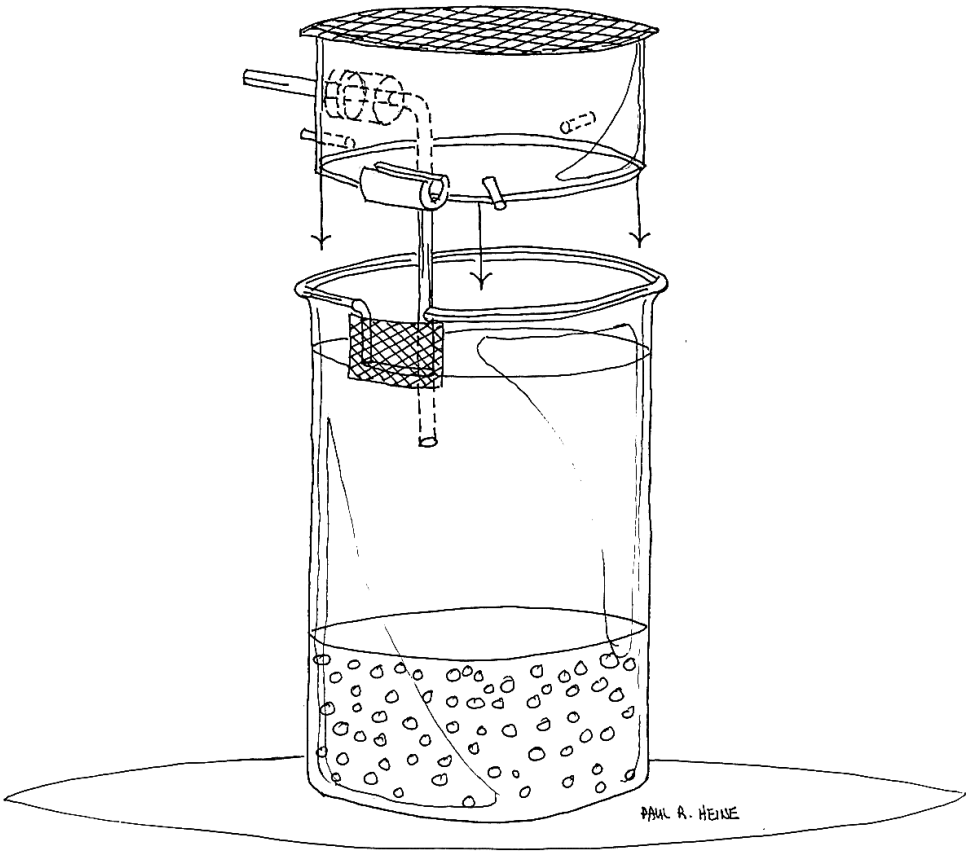


Figure C.4 Emergence traps that can be used with the Zumwalt water-delivery system described in Section A.4.

to be aerated during the midge exposure. An air line is connected to each tube and a gang valve is used to regulate air flow (about 1 bubble/second). The glass tube extends below the bottom of the plexiglass tube into the surface of the overlying water. A 4-mm slot will need to be cut in the petri dish in order to slide the petri dish under the emergence trap to remove adult midges from the test beakers (Figure C.2). The emergence trap capped with this petri dish can then be set on a 300-mL beaker to remove the adults with an aspirator as illustrated in Figure C.3.

C.9.7 Press 3/8-inch-long pins into the three 1/16-inch holes drilled in the side of the plexiglass tube. These pins make the plexiglass tube stable on the top of the beaker.

C.9.8 If the plexiglass tubes are used in beakers with a notch at the top (i.e., the beakers described in Zumwalt et al., 1994; Section A.4), a 2-cm length of 1/8-inch inner diameter latex tubing will need to be slit lengthwise and then slipped onto the bottom of the plexiglass tube. This tubing is then lined up with the notch in the beakers to prevent emerging midges from escaping. This piece of tubing is not needed if beakers described in Benoit et al. (1993) are used (i.e., beakers with holes drilled in the side).

C.9.9 Supplies

A. McMaster Carr, P.O. Box 4355, Chicago, IL 60680-4355, 708/833-0300 (part number and materials).

1. 8486 K 115, Acrylic tube 2 1/2-inch outer diameter and 1/8-inch wall.

2. 9226 T 84, 16- X 16-inch stainless wire cloth (0.018-wire diameter).

3. 90145 A 417, 1/16-inch diameter stainless dowel pins 3/8 inch long.

B. Thomas Scientific, P.O. Box 99, Swedesboro,NJ 08085-0099, 609/467-2000: 8747-E17, #00 silicone stopper.

Appendix D
Sample Data Sheets

Culture Aquarium	Date of Egg Mass Deposition	Date 4th Instar Larvae Were Weighed	Age of Weighed 4th Instar Larvae	Mean Dry Weight of 4th Instar Larvae (n = 10)	Date of Observed First Emergent Adult	Total Number of Egg Masses Produced	General Comments	Initials of Culturist
A								
B								
C								
D								
E								
F								

Figure D.1 Data sheet for the evaluation of a *Chironomus tentans* culture.

Position # _____. Tank # _____. Set up Date ___ / ___ / ___. Init. _____.

Embryo Deposition Date ___/ ___ / ___.
Embryo Hatch Date (day 0) ___/ ___/ ___.
Number of larvae used to initiate tank _____.
or number of egg cases used _____.
Date 10 Days Old Post Hatch ___/ ___/ ___.
First Emergence Date ___/ ___ / ___.
Substrate Type _____.
Food Type _____. Conc. _____. Date Made _____.

Emergence Data (Performed 3 x Per Week)

Date	# of	# of	Total	Comments	Init.

Chemistries (Performed Weekly)

Date	pH	D.O.	Ammonia	Temperature	Init.

Figure D.2 QA/QC data sheet for *Chironomus tentans* culture.

10-d Old *C. tentans* Body Lengths		
Date:		
Tank #	Length	Init.
Mean		

10-d Old *C. tentans* Head Capsule Widths		
Date:		
Tank #	Width	Init.
% 2nd Instar		
% 3rd Instar		
% 4th Instar		

***C. tentans* Dry Weight Data**						
Date	Tank #	Pan + 10 Organisms	Pan Only	Difference	Weight/ Organism (mg)	Init.
Mean						

Figure D.3 QA/QC data sheets for *Chironomus tentans* culture.

Brood Stock Source_____

Test Type (circle one)[1]: SU SM RU RM FU FM

No. of Animals Tested Per Replicate_____

No. of Replicates_____

Method of LC50 Estimate_____

Reference Toxicant ($CuSO_4$ or KCl)_____

Reference Toxicant Supplier and Lot No.

_____

Reference Toxicant Purity_____

Test Initiation Date_____

Toxicologist_____

Exposure Duration (Hr)	Number of Mortalities											
	Control		Exp. 1		Exp. 2		Exp. 3		Exp. 4		Exp. 5	
	A	B	A	B	A	B	A	B	A	B	A	B
0												
24												
48												
72												
96												

Current Test 96-h LC50 = _____

Number of Reference-toxicity Test Used
to Determine Cumulative Mean 96-h LC50_____

Mean 96-h LC50 for All Tests to Date_____

Acceptability of Current Test[2] Yes_____ No_____

---

[1] SU = Static unmeasured
SM = Static measured
RU = Renewal unmeasured
RM = Renewal measured
FU = Flow-through unmeasured
FM = Flow-through measured

[2] Based on two standard deviations around the cumulative mean 96 h-LC50

Figure D.4 Data sheet for performing reference-toxicity tests.

Sediment Sample Source_____

Date of Test Initiation_____

Toxicologist Conducting Test_____

Test Day	Test Replicate Sampled	Temperature (^{o}C)	Dissolved Oxygen (mg/L)	pH	Hardness (mg/L)	Alkalinity (mg/L)	Specific Conductance (umhos/cm)	Total Ammonia (mg/L)
0								
1								
2								
3								
4								
5								
6								
7								
8								
9								
10								

Figure D.5 Data sheet for temperature and overlying water chemistry measurements.

Daily Checklist for Sediment Tests

Dissolved Oxygen
Minimum Acceptable Concentration
(40% of Saturation at Target Temp)

Study Code _____

Study Name _____

Building _____ Diluter _____ Waterbath _____ Target temperature _____ °C

Study Director _____ Acceptable Range _____ °C to _____ °C

Lead Technician _____ Month _____ = _____ mg/L

Day of Month	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
Day of Study																															
Diluter Operation																															
Number of Cycles																															
Time of Day																															
Temperature																															
Air Pressure																															
Aeration																															
Brush Screens																															
Clean Needles																															
Feeding																															
Total Water Quality																															
Partial Water Quality																															
Initials																															

Comments _____

Approved by _____ Date _____

Figure D.6 Data sheet for daily checklist for sediment tests.

Water Quality Data Sheet

Study Code _____ Date _____ Test Day _____ Study Director _____ Investigator _____

Study Name _____

Sample Code	Dissolved Oxygen (mg/L)	Temp. °C	Salinity (ppt)	Conductivity (umhos or uS)	pH	Alkalinity (ppm $CaCO_3$) (ml Titrant x mult. factor =)			Alkalinity (ppm $CaCO_3$) (ml Titrant x mult. factor =)			Hardness (ppm $CaCO_3$) ___ml	Turbidity (NTU) ___ml	Total Ammonia as NH_3 (mg/L) ___ml	ml of ISA	Other _____ ___ml	Other _____ ___ml
Volume of Sample =	___ml	___ml	___ml	___ml	___ml		___ml										
_____	_____	_____	_____	_____	_____	_____ x	_____	= _____	_____ x	_____	= _____	_____	_____	_____	_____	_____	_____
_____	_____	_____	_____	_____	_____	_____ x	_____	= _____	_____ x	_____	= _____	_____	_____	_____	_____	_____	_____
_____	_____	_____	_____	_____	_____	_____ x	_____	= _____	_____ x	_____	= _____	_____	_____	_____	_____	_____	_____
_____	_____	_____	_____	_____	_____	_____ x	_____	= _____	_____ x	_____	= _____	_____	_____	_____	_____	_____	_____
_____	_____	_____	_____	_____	_____	_____ x	_____	= _____	_____ x	_____	= _____	_____	_____	_____	_____	_____	_____
_____	_____	_____	_____	_____	_____	_____ x	_____	= _____	_____ x	_____	= _____	_____	_____	_____	_____	_____	_____
_____	_____	_____	_____	_____	_____	_____ x	_____	= _____	_____ x	_____	= _____	_____	_____	_____	_____	_____	_____
_____	_____	_____	_____	_____	_____	_____ x	_____	= _____	_____ x	_____	= _____	_____	_____	_____	_____	_____	_____
_____	_____	_____	_____	_____	_____	_____ x	_____	= _____	_____ x	_____	= _____	_____	_____	_____	_____	_____	_____
_____	_____	_____	_____	_____	_____	_____ x	_____	= _____	_____ x	_____	= _____	_____	_____	_____	_____	_____	_____
_____	_____	_____	_____	_____	_____	_____ x	_____	= _____	_____ x	_____	= _____	_____	_____	_____	_____	_____	_____
Meter #	_____	_____	_____	_____	_____			_____				_____	_____	_____	_____	_____	_____
Initials	_____	_____	_____	_____	_____		_____					_____	_____	_____	_____	_____	_____

Comments _____

Approved by _____ Date _____

Figure D.7 Data sheet for water quality parameters.

CHEMISTRIES

Page ____ of ____

Test Type _____ Sample Info _____ Water Type _____

Organism _____ _____ Experimenter _____

Test Dates _____ _____ Test System _____

I.D.	Day	-1	0	1	2	3	4	5	6	7	8	9	10	Remarks
pH														
DO (mg/L)														
Temp ^{0}C														
Hard/Alk														
pH														
DO (mg/L)														
Temp ^{0}C														
Hard/Alk														
pH														
DO (mg/L)														
Temp ^{0}C														
Hard/Alk														
pH														
DO (mg/L)														
Temp ^{0}C														
Hard/Alk														

Figure D.8 Chemistry data sheet.

Study Director _____

Study Code _____

Study Name_____

Daily Comment Sheet

Day_____ Date_____-_____-_____ Initials_____

Day_____ Date_____-_____-_____ Initials_____

Day_____ Date_____-_____-_____ Initials_____

Day_____ Date_____-_____-_____ Initials_____

Day_____ Date_____-_____-_____ Initials_____

Figure D.9 Daily comment data sheet.

Weight Data Form

Test Dates			Species				
Test Material			Weighing Date			Food	
Location			Oven Temp (°C)			Age Organisms	
Analyst			Drying Time (h)			Initial No/Rep	
Sample	Replicate	Wt. of Oven Dried Pan (mg)	Wt. of Pan + Oven Dried Organisms (mg)	Dried Wt. of Organisms (mg)	Number of Survivors	Mean wt per Survivor	Sample Mean

Figure D.10 Weight data sheet.

Date: _____ **Test:** _____

Species: _____ **Investigator:** _____

Facility: _____

Treatment (Site)	Rep	Number Surviving	Pan Weight	Pan + Larvae	Dry Weight Total	Dry Weight Indiv.	Pan + Ash	Ash-free Dry Wt Total	Ash-free Dry Wt Indiv.
	1								
	2								
	3								
	4								
	5								
	6								
	7								
	8								
	1								
	2								
	3								
	4								
	5								
	6								
	7								
	8								
	1								
	2								
	3								
	4								
	5								
	6								
	7								
	8								

Figure D.11 Data sheets for *Chironomus tentans* tests.

At termination of test:

1. Sieve sediment from each beaker and record the number of recovered larvae in the "survival" column.

2. Place all larvae from one replicate in a pre-ashed and pre-weighed aluminum weigh pan.

3. Dry larvae at 60°C for at least 24 hr.

4. Weigh pan + larvae and record weight under appropriate column of data sheet.

5. Ash pan + larvae at 550°C for 2 hr. Let cool to room temperature.

6. Weigh pan + ashed material.

7. Remove ash (e.g. with a small brush) and weigh pan.

8. Calculate dry weight as the difference between the pan+larvae weight and the pan weight.

9. Calculate ash-free dry weight as the difference between the pan+larvae weight and the pan+ash weight.

Figure D.12 Instructions for terminating a *Chironomus tentans* test.

Date: _____ Test: _____ Species: _____

Investigator: _____ Facility: _____

Treatment	Rep	Larvae Number	Dead Pupae			Date of Emergence				Date of Egg Mass	Egg Counts		Number Eggs Not Hatched	Date Adult Died
						Partial		Complete						
			No ID	♂	♀	♂	♀	♂	♀		Live	Acid		
		1												
		2												
		3												
		4												
		5												
		6												
		7												
		8												
		9												
		10												
		11												
		12												

Comments (Adult transfers, mate pairings etc.)

Data Summary

No. of larvae recovered at end of test:	Total Larvae:	Number Dead Pupae:
No. Dead/Escaped Adults:	Total Emerged Adults:	Total Egg Mass:

Figure D.13 Data sheet for the *Chironomus tentans* life-cycle test.

Date: 01/28/96 Test: Mississippi River Species: *C. tentans*

Investigator: _____ Facility: _____

	Rep	Larvae Number	Dead Pupae			Date of Emergence				Date of Egg Mass	Egg Counts		Number Eggs Not Hatched	Date Adult Died
			No ID	♂	♀	Partial ♂	Partial ♀	Complete ♂	Complete ♀		Live	Acid		
7RR-A	A	1						2/23[a]						2/23
		2						2/24[b]						
		3							2/26	2/28		644		3/3
		4						2/25						3/2
		5						2/27						
		6							2/27	2/28	722		38	3/4
		7						2/27[c]						3/6
		8	3/4											
		9						3/6[d]						
		10							3/8[e]	3/9	1004		31	3/13
		11	3/15											
		12												

Comments (Adult transfers, mate pairings etc.)

[a] Fully emerged; dead on water	[e] 3/8 ♂ transferred from 7SR-A on 3/8	
[b] 2/24 ♂ transferred to 7RR-B on 2/24		
[c] 2/27 ♂ transferred to 7RR-B on 3/2		
[d] 3/6 ♂ transferred to 7RR-C on 3/6		

Data Summary

No. of larvae recovered at end of test: 0	**Total Larvae: 11**	**Number Dead Pupae: 2**
No. Dead/Escaped Adults: 1/0	**Total Emerged Adults: 9**	**Total Egg Masses: 3**

Figure D.14 Example entries for a *Chironomus tentans* life-cycle test data sheet.

Copy of a sample data sheet that will be used to record all information pertaining to emergence and reproduction of *C. tentans* during the life-cycle test. For clarity, consistency, and ease of data interpretation, it is important that each lab fill out this sheet as illustrated. A brief interpretation of each recording (column) is provided below.

I Data Sheet Requirement. One data sheet is needed for each replicate. Thus, a treatment having 8 reproduction replicates will have 8 data sheets (survival and growth data are recorded on separate sheets). All emergence and reproduction data for a replicate are recorded on the corresponding data sheet.

II Recording Pupae, Emergence, and Egg Mass Data. Record all pupae, emergence, and egg mass data as dates.

III Column Heading Interpretation

Station/Site and Replicate. Enter name of sample and corresponding replicate (e.g., 7RR-A).

Larvae #. These numbers correspond to the 12 larvae placed in each replicate.

Dead Pupae. If it is not possible to determine the gender of the dead pupae, enter the date found in the **"No ID"** column. Otherwise, enter the date found in either the male or female column.

Date of Emergence. If an adult has not completely shed the pupal exuviae, enter the date found under the **"partial emergence"** category as a male or female. If emergence is complete but the adult is dead (typically floating on the water surface), record date under **"complete emergence"** category as a male or female and enter a footnote as indicated in **"footnote a"** in comments section of data sheet.

Partially emerged adults, and those that have emerged completely but were unable to escape the surface tension of the water, usually die within 24 hr. In both cases, the date of death should be recorded as one day later under the **"Date Adult Died"** column.

Date of Egg Mass. Record the date on which the egg mass was collected from the replicate.

Egg Counts. Enter number of eggs counted using either the acid-digestion (direct count) or ring method (indirect count).

Number Eggs Not Hatched. Enter the number of unhatched eggs from each oviposited egg mass for which an indirect count (ring method) was determined.

Date Adult Died. Enter the date that the adult died (be sure to follow transferred adults).

IV Comments Section. All comments concerning adult transfers and emergence patterns should be recorded in this section as footnotes (see footnotes a-e on sample data sheet).

V Data Summary Section. At termination of each replicate, record the **Number of Larvae Recovered at End of Test** after sieving and determine the number of **Total Larvae** alive during the test. Also record the **Number Dead Pupae**, **Number Dead/Escaped Adults, Total Emerged Adults,** and number of **Total Egg Masses** by summing the appropriate columns.

VI Example Entries for *C. tentans* Data-Sheet 7RR-A

Example #1. On 2/23/95 a male emerged from this replicate. This is recorded under the "Male" category of the "Complete Emergence" column on the first line. This male was fully emerged but was dead and floating on the water surface. This is recorded as footnote "a" in the "Comments" section and the date of death recorded under the "Date Adult Died" column.

Example #2. A female emerged from this replicate on 2/26/95 which is recorded under the "Female" heading of the "Complete Emergence" column. This female produced an egg mass on 2/28/96 which is recorded under the "Date of Egg Mass" column.

Example #3. A dead pupae was recorded on 3/4/95. Since the sex was not determined, it was recorded under the "No ID" heading of the "Dead Pupae" column. Pupal sex may be determined by examining the genitalia under a dissecting microscope (the genitalia can be seen through the pupal exuviae which is usually, but not always, transparent).

Example #4. A male emerged on 2/24/95 in 7 RR-A and was transferred to replicate 7RR-B. This is shown as footnote "b". Recording this type of data helps to keep track of where males are and the number of times they have reproduced.

A male from 7SR-A (one of the stand-by replicates) was transferred to 7 RR-A on 3/8/95. This is recorded as footnote "e" on the 7RR-A data sheet. For completeness, a corresponding footnote on the 7 SR-A data sheet should be made regarding this transfer.

D.15 Instructions for completing the *Chironomus tentans* life-cycle test data sheet.