Laboratory Guidance and Whole Effluent Toxicity Test Review Criteria

Publication No. WQ-R-95-80
Revised December 1998

Printed on Recycled Paper

Laboratory Guidance and Whole Effluent Toxicity Test Review Criteria

Prepared by:

Washington State Department of Ecology
Water Quality Program
12/17/98

Publication No. WQ -R-95-80
Revised December 1998

Printed on Recycled Paper

For additional copies of this document, contact:

Department of Ecology
Publications Distribution Center
PO Box 47600
Olympia, WA 98504-7600
Telephone: (360) 407-7472

Table of Contents

I. Whole Effluent Toxicity Testing Regulatory Guidance

A. Introduction

On November 6, 1993, a new rule became effective in the state of Washington: Chapter 173-205 WAC Whole Effluent Toxicity Testing and Limits. The short name for this rule is the whole effluent toxicity (WET) rule. Chapter I. Whole Effluent Toxicity Testing Regulatory Guidance of this document has been prepared to assist labs in providing toxicity testing services to permittees who must meet the requirements of the WET rule. The guidance will help provide the regulatory context for WET testing and other services provided by labs. Having an understanding of the purpose of WET testing can help labs provide better service to permittees.

Chapter II. Whole Effluent Toxicity Test Review of this document has been prepared to assist accredited labs to provide acceptable toxicity tests for permittees who are regulated under the WET rule. Only WET tests and rapid screening tests from accredited labs can be used to fulfill these requirements.

A National Pollutant Discharge Elimination System (NPDES) permit will describe which requirements of the WET rule apply to each individual permittee and what specific actions the permittee must take to meet these requirements. An administrative order can also be used to communicate these requirements. This document does not supersede or modify the requirements of any valid permit unless the permit references an outdated test manual. If this document seems to conflict with the requirements in a permit, it is likely that the permit was written before the WET rule or this guidance was written. These older permits, including expired permits, are still valid permits. If a lab believes that any permit requirement could be improved by making it more consistent with this document, then the permittee can be advised to contact the Department of Ecology (Ecology) to request a change. (See WAC 173-205-080(1)(c).) However, labs should not deviate from the instructions in any valid permit unless the deviation has been approved by Ecology.

All questions concerning this document or the WET testing program should be directed to Randall Marshall (360-407-6445) or Keith Johnson (360-407-6442).

B. WET Testing Requirements in NPDES Permits

Effluent Characterization

Effluent characterizations last for one year. During this year, each effluent sample is tested with all of the WET test species listed in the permit. This "multiple species" testing provides an assessment of the toxicity of the effluent sample to different types of aquatic organisms.

Effluent characterization is used to establish whether a WET limit is required. After effluent characterization, a permittee might receive an acute WET limit, a chronic WET limit, both WET

limits, or no WET limit. Permittees who cannot meet the WET performance standards defined in the WET rule will receive WET limits.

For acute toxicity, the performance standard is a median of 80 percent survival in 100 percent effluent at the end of effluent characterization with no single test result showing less than 65 percent survival in 100 percent effluent.

For chronic toxicity, the performance standard is no statistically significant difference in test organism response between the control and a test concentration equal to the concentration of effluent at the edge of the acute mixing zone (acute critical effluent concentration or ACEC).

If a mixing zone has not been established for the discharge at the time of permit writing, the ACEC will not be known during effluent characterization. When the ACEC is unknown, WET testing during effluent characterization will determine the no observed effect concentration (NOEC). The NOECs will be compared to the ACEC, when it becomes known, to determine if a chronic WET limit is needed. If the ACEC is still unknown at the end of effluent characterization, then effluent characterization will be extended, but only one WET test will be conducted on each sample ("single species" testing).

It is in the permittee's best interest to include the ACEC in the dilution series as soon as it becomes known because the permittee will be at a disadvantage whenever the ACEC would have been between the LOEC and NOEC.

Effluent characterization is also used to establish a baseline toxicity level expressed by point estimates such as the LC_{50}, EC_{50}, or IC_{25}. These point estimates will not be used in determining compliance, but will serve as a point of reference if problems with toxicity need to be investigated. WET tests conducted for effluent characterization must have a dilution series of at least five effluent concentrations in order to provide point estimates.

Compliance Monitoring

The state's Water Quality Standards prohibit toxicity past the edge of an approved mixing zone. Therefore, WET limits are based on the concentration of effluent at the edge of an approved mixing zone during critical conditions. Critical conditions are situations when the effect of the effluent is greatest such as during low river flow. The concentration of effluent existing at the edge of a mixing zone during critical conditions is called the critical effluent concentration. Compliance with a WET limit means demonstrating no toxicity in a sample of effluent diluted to equal the critical effluent concentration. The ACEC used to test for compliance with an acute WET limit (and as the chronic performance standard as described above) is the concentration of effluent at the edge of the acute mixing zone. The chronic critical effluent concentration (CCEC) used to test for compliance with a chronic WET limit is the concentration of effluent at the edge of the chronic mixing zone.

A permittee complies with a WET limit when the hypothesis testing procedure in Appendix H of EPA/600/4-89/001 (Fisher's Exact Test for survival in the *Ceriodaphnia* chronic test) has shown no statistically significant difference in response between the ACEC or CCEC and a control. Appendix H of EPA/600/4-89/001 is the same as Appendix H in the new freshwater chronic manual and Appendix G in the new marine chronic manuals. The new EPA acute manual

describes the single comparison hypothesis testing procedure on pages 101-105. A statistically significant difference in test organism response (alpha = 0.05) would mean a WET limit violation. (See Appendix D, Identifying Anomalous WET Tests, for exceptions to this.)

WET testing to monitor for compliance with an acute WET limit must be conducted at a minimum with the ACEC (the limit), 100 percent effluent (the performance standard), and a control. The permittee may request a full dilution series to provide more information for review of test quality.

WET testing to monitor for compliance with a chronic WET limit must be conducted with the CCEC (the limit), the ACEC (the performance standard), and a control. The permittee may request a full dilution series to provide more information for review of test quality.

Monitoring for Changes in Toxicity

Permittees not given WET limits after effluent characterization will not be conducting compliance monitoring for WET. However, the WET rule does require these permittees to demonstrate that toxicity has not increased during the permit term. If toxicity has increased, then a new effluent characterization will be required. The WET rule specifies several types of actions that permittees might make in order to demonstrate that toxicity has not increased. These actions include:

- The WET Rule allows Ecology to condition the non-assignment of a WET limit on routine monitoring with a rapid screening test if there is the potential for an event at the facility which could result in a toxic discharge that would otherwise go unnoticed.

  A rapid screening test is a single dilution (plus a control) toxicity test on 100 percent effluent or the ACEC in order to detect unanticipated increases in toxicity. Rapid screening tests are less expensive and quicker than the standard WET tests used for effluent characterization or compliance monitoring. (See Appendix F for the list of rapid screening tests.)

  Whenever a permittee fails a rapid screening test, the WET rule requires the permittee to immediately retest with standard WET tests. The results of these WET tests conducted in response to rapid screening tests will be evaluated to determine the need for a new WET characterization in the next permit or the need for administrative orders to immediately investigate and control toxicity. Compliance with WET limits will not be measured with rapid screening tests.

- The WET rule requires that permittees without a WET limit who are not conducting rapid screening testing must submit a set of WET test results with each permit application. These WET tests would be the same standard WET tests used in effluent characterization. In most cases, Ecology would require only a few WET tests be conducted for submission with the permit application. However, the set of WET tests required for permit application would be larger if any of the WET tests conducted for effluent characterization was unacceptable (See Chapter II. Whole Effluent Toxicity Test Review and Appendix D Identifying Anomalous WET Tests.) and Ecology needed additional WET test results to complete the effluent characterization.

- The WET rule requires permittees to evaluate any changes with the potential to increase effluent toxicity. Compliance monitoring or rapid screening testing are assumed to accomplish this evaluation automatically. For other permittees without WET limits or rapid screening testing, extra WET tests may have to be conducted when a change occurs at the facility although other techniques, such as chemical analysis, may be employed to demonstrate that toxicity has not increased.

Response to Noncompliance with a WET Limit

If a permittee fails a compliance test for a WET limit, then additional testing is immediately required to assess and confirm the continuing presence of toxicity. The WET Rule requires WET testing of four weekly samples following noncompliance with an acute WET limit and three monthly samples following noncompliance with a chronic WET limit. If any of these additional WET tests fails to comply with a limit, then the permittee must submit a toxicity identification/reduction evaluation (TI/RE) plan.

Permit Language

New permit language for WET requirements can be complicated. Permit language will contain a series of steps in a regulatory process. The step to follow will depend at times on the results of the previous step. The permit might contain two sets of instructions, but only require that one set be followed depending on circumstances. This permit language prevents the extra expense and effort associated with permit modifications, but will require careful reading and planning ahead by labs and permittees.

Researching Specific Problems

A problem such as the failed smoltification of salmon in the vicinity of an outfall might be researched using WET testing. However, it is likely that the WET rule would not allow such testing to be used for effluent characterization or compliance monitoring, and it would have to be evaluated outside of the context of the WET rule.

C. Options for Permittees

The WET rule contains options for permittees to use if they decide that it is in their best interest to do so.

Full Dilution Series Tests

WET tests conducted using a full dilution series of at least five effluent concentrations and a control provide the best information for evaluating the quality of WET test results. A full dilution series protects permittees by allowing anomalous test results to be identified more easily. Anomalous WET tests will not be used for compliance determinations. Because the WET rule allows WET tests in some circumstances to be conducted with less than a full dilution series, it also makes clear that permittees may choose to conduct any WET test using a full dilution series. The ACEC or CCEC may be included in any dilution series as an extra concentration or as a substitute for a standard concentration in the series.

Effluent Screening Tests

The WET rule allows Ecology to approve the request of a small business or the request of a POTW discharging less than 0.5 mgd to conduct WET testing using effluent screening tests. Effluent screening tests are WET tests that are conducted using only a control and 100 percent effluent for an acute WET test or only a control and the ACEC for a chronic WET test. If the effluent screening test shows toxicity, the permittee is required to resample and conduct a full dilution series WET test.

Sample Handling and Testing Requirements not in Accordance with the WET Rule

The WET rule contains instructions for some aspects of sample handling and toxicity testing such as when dechlorination is acceptable, which test methods are approved, and the duration of acute tests. New permits will contain instructions that meet these requirements of the WET rule. Some older permits might contain requirements that conflict with the WET rule. (See Chapter II. Whole Effluent Toxicity Test Review and Chapter III. Toxicity Test Report Checklist.]

The prompt replacement of any inappropriate sample handling or toxicity testing requirement will minimize the need to conduct additional toxicity tests in order to provide an adequate effluent characterization. WAC 173-205-080(1)(c) allows Ecology to approve the request of any permittee whose permit predates the WET rule to replace inappropriate requirements with appropriate ones. Even though labs have no requirement to do so, they are particularly well-placed to identify and inform permittees of testing requirements that need to be changed.

Notification of an Anomalous Test Result

The WET rule allows a permittee to avoid the cost of additional testing when noncompliance with a WET limit is believed to be due to an anomalous WET test result. A laboratory should be able to inform a permittee of any anomalous WET test result that resulted in noncompliance with a WET limit. (See Appendix D, Identifying Anomalous WET Tests.) The permittee then sends Ecology notification with the compliance test report that the test might be anomalous and that the permittee intends to take only one additional sample for toxicity testing. The notification must identify the reason for considering the compliance test result to be anomalous. If Ecology agrees that the test causing noncompliance was anomalous, then the permittee is saved the cost of the rest of the additional testing. The one additional test will replace the anomalous test.

II. Whole Effluent Toxicity Test Review

A. Introduction

On November 6, 1993, a new rule became effective in the state of Washington: Chapter 173-205 WAC Whole Effluent Toxicity Testing and Limits. The short name for this rule is the whole effluent toxicity (WET) rule. Chapter II. Whole Effluent Toxicity Test Review of this document has been prepared to assist accredited labs to provide acceptable toxicity tests for permittees who are regulated under the WET rule. Only WET tests and rapid screening tests from accredited labs can be used to fulfill these requirements.

A National Pollutant Discharge Elimination System (NPDES) permit will describe which requirements of the WET rule apply to each individual permittee and what specific actions the permittee must take to meet these requirements. This document does not supersede or modify the requirements of any valid permit unless the permit references an outdated test manual. If a lab believes that any permit requirement could be improved by making it more consistent with this document, then the permittee can be advised to contact the Department of Ecology (Ecology) to request a change. (See WAC 173-205-080(1)(c).) However, labs should not deviate from the instructions in any valid permit unless the deviation has been approved by Ecology.

The test review criteria and appendices in this document have been reviewed and commented on by the accredited labs and other interested parties. The document was revised in response to persuasive comments given by labs.

Questions concerning this document or the WET testing program should be directed to Randall Marshall (360-407-6445) or Keith Johnson (360-407-6442).

B. Invalid Tests

Invalid WET tests occur when the lab does not follow the test method or when the results do not meet the validation criteria in the test method. Permittees are obligated to look for invalid tests because the permit requires that only the results of valid tests be submitted. Ecology will review WET test results to see that they are based on valid tests. In addition to the items in this section, the EPA manuals and Chapter III. Toxicity Test Report Checklist will be used to test validity.

1. Failure of EPA Statistical Flowcharts

A WET test is considered invalid and must be repeated if the flowcharts for determining NOECs in the EPA toxicity test manuals cannot be followed due to a low number of replicates. The problem will occur when there are less than four replicates and the test data are not normally distributed or have unequal variances. The number of replicates is more important in hypothesis testing than in point estimations, and the minimum number of replicates in the EPA manuals is sometimes too low for determining NOECs correctly even when point estimation works fine. Labs should be aware of the EPA recommendation to use the Kolmogorov “D” statistic to replace Shapiro-Wilk’s test when $n > 50$. [The flow

chart for the process (single comparison hypothesis testing) in Appendix H of the EPA freshwater chronic manual and Appendix G of the marine chronic manuals can be found in Figure 12 of the acute manual, EPA/600/4-90/027F. This flowchart must also be successfully followed.]

Four replicates will often be inadequate for determining an NOEC when replicate numbers are unequal and test data are not normally distributed or have unequal variances. Labs intending to run extra control replicates should consult the table of critical values for Wilcoxon's Rank Sum test to determine the minimum number of replicates at the test concentrations. The accidental loss of a test chamber in a typical test of five test concentrations and a control will also cause replicate numbers to be unequal and four replicates to be inadequate if the nonparametric hypothesis test (Wilcoxon's Rank Sum) must be used. The minimum number of replicates required will not be increased for those tests where it currently stands at four because the accidental loss of test chambers is not a frequent occurrence and will rarely necessitate the rejection of a test for failing the EPA statistical flowchart. If a test chamber has been accidentally lost from a test using four replicates/concentration and requiring a nonparametric hypothesis test, then the concentration-response relationship will be examined to see if the concentration losing a replicate can be excluded from the analysis because it appears to be nontoxic (healthy test organism performance nearly equal to adjacent concentrations and the control) or if it and adjacent concentrations have a nearly complete adverse effect (complete mortality, loss of neonate production, etc.). If the ACEC and CCEC have been included in the concentration series of a test losing a test chamber and have at least three replicates remaining at the end of the test, then single comparison hypothesis testing can be used to compare the ACEC or CCEC to the control.

If a lab increases the number of effluent concentrations in a test series beyond five, the EPA flowcharts for determining NOECs may not work. Adding extra concentrations to the series improves the ability of a test to measure toxicity and calculate point estimates. Unfortunately, the extra concentrations also raise the minimum number of replicates required for determining an NOEC to five or higher under some circumstances (such as Steel's many-one rank test and Wilcoxon's rank sum test).

Assuming that at least four replicates were used, a test with more than five effluent concentrations in the series is still valid even when the EPA flowchart for determining an NOEC fails. Removing one or more of the concentrations from the series before attempting to determine the NOEC will solve the problem without having to increase the number of replicates beyond four. All effluent concentrations in the test should be used to calculate point estimates and be included in the test report, but it is acceptable to exclude one or two concentrations from the NOEC determination in order to successfully follow the EPA flowchart. The concentrations that are removed from consideration should be as far from the threshold of toxic response (LOEC/NOEC) as possible.

An important point to note on this subject is that labs are free to perform statistics in any way they feel is appropriate to meet the client’s needs and to report results accordingly. When we review the test results, we will recalculate the statistics as described in this document and the permit and will insist only that the test be conducted (number of

replicates, etc.) and data recorded so that we can successfully perform the statistics. Our decisions will be based on our own calculations.

2. Appropriate Negative Controls

Negative controls serve two important functions in toxicity tests:

- **Establishing test validity -** A control provides a measure of test organism health and laboratory technique in order to establish the validity of the test result. Every toxicity test must have a control that accomplishes this function. For acute toxicity tests conducted during effluent characterization, this is the primary function for the control because no hypothesis testing is needed.

- **Providing a standard for comparison in hypothesis testing -** The control in a valid toxicity test also provides an indication of test organism response under nontoxic conditions. The control response can then be compared to organism response in an effluent concentration using hypothesis testing in order to determine if the effluent is toxic at that concentration.

To accomplish these functions, it is important that controls are nontoxic laboratory or natural water, that the same water is used for both the control and diluting the sample, and that controls are handled the same as all other test concentrations. A toxicity test is not acceptable unless the control meets these conditions.

In order to use one control in testing more than one sample, a lab must demonstrate in the standard operating procedure (SOP) approved as a part of accrediting the lab for the test method that all of these important conditions are being met. The randomization of the control with test containers from all samples is especially important (See the first paragraph in Appendix A of any of the EPA toxicity test manuals listed at the bottom of page 12). Every test container for every sample sharing a control should be handled as if part of one large test with all activities occurring within the same space and time. Implementation of the procedure must also be documented for all tests sharing one control. Failure to do so will cause test results to be rejected.

One misuse of a control which will certainly result in rejection of the toxicity test result is running extra replicates in the control and only using the results from the replicates with the best performance. Controls must be handled the same as other test concentrations. Failure to do so will cause rejection of the test.

3. Appropriate Test Termination

All tests must be continued for the full duration specified in the permit or test protocol. If all test organisms die in every test concentration, the control must still be continued for the full duration in order to produce acceptable test results. It is acceptable to terminate a test early which, if continued, would not meet the requirements of the permit or test protocol as long as the effluent is resampled immediately and an acceptable test result produced as soon as possible. An explanation of the reasons for early termination must accompany the report for the test on the new sample.

4. Acceptable Start Counts

The EPA statistics are based on the assumption of equal numbers of test organisms in each replicate at the start of a test. Small deviations (one or two test organisms) from equality will not cause a problem with statistics, but larger differences will put the validity of statistics in doubt. Labs should not 'cut corners' by not properly recounting the number of organisms in each replicate immediately after test initiation. Start counts may be changed based on the discovery of a miscount during the recount immediately after test initiation but not afterwards unless due to albino fathead minnows.

The loss of controlled experimental conditions is important in evaluating test validity when the number of test organisms was not equal in the replicates at the beginning of the test. If the number of organisms in the replicate containers is unequal, then either the amount of food/animal must be unequal or the amount of food/test solution volume must be unequal. If the number of organisms in the replicate containers is unequal, then either the test organism loading must be unequal or the test solution volume must be unequal. Unequal numbers of test organisms in replicates will always create other inequalities of test conditions. The integrity of the test design is compromised.

Toxicity tests with large or frequent differences in test organism numbers in the replicates will be rejected and returned to the permittee. Toxicity tests run on future samples will be rejected if the organism start count is not equal in the replicates. No more than three replicates out of 24 (approximately 10 percent) can vary in organism start count in any individual test or the test will be rejected. No more than 10% of the toxicity tests conducted by any one lab in a year should vary in start count or permittees will be notified.

If test organisms are lost or killed by a documented accident, then the start count should be appropriately reduced. Accidents are specific events usually caused and observed by people. Examples of accidents include spilling, siphoning, or crushing test organisms. If aeration is necessary in order to maintain adequate dissolved oxygen during a test, then any test organisms found stranded on the side of the test chamber, caught in the test solution's surface tension, or entrained in an air bubble can be assumed to be victims of an accident. Test organism cannibalism, stranding on the side of the test chamber (unless due to aeration or agitation of the test chamber during handling), or simple disappearance are not documentable accidents and do not justify adjusting start counts. Test organism weakness or death often precedes cannibalism, stranding, or disappearance. Cannibalism should be controlled by generous feeding (but not significantly overfeeding), and stranding can be minimized by avoiding supersaturation or excessive shaking of test chambers. Tidying-up the data by adjusting start counts and thereby reducing variation is especially unfair when hypothesis testing is used to make regulatory decisions. The limit on varying start counts mentioned above applies also to adjusted start counts.

5. Acceptable pH Adjustment

If the sample pH is outside of the range 6.0 to 9.0, then the permittee is likely to be in violation of a technology-based permit limit for pH and could also be violating water quality standards. Permittees should be immediately alerted to a potential problem if this occurs. Samples outside of this range will be rare.

Labs are forbidden from adding acids and bases to samples because manipulation of samples (aeration, filtration, addition of acids, bases, or sodium thiosulfate, etc.) should be minimized. In principle, no substance should be introduced into the sample unless absolutely necessary for a successful toxicity test. Acids and bases might themselves be toxic or enhance the toxicity of other substances.

Every effluent sample must be tested without pH adjustment regardless of initial pH. Labs may adjust the pH of a portion of a sample which is outside of the 6.0 to 9.0 pH range to pH 7.0 for freshwater testing or pH 8.0 for saltwater testing. If pH adjustment is done, the test must be conducted in parallel with a portion at one or more concentrations pH adjusted, and a full test run without adjustment for the entire concentration series.

Parallel testing of pH adjusted and unadjusted sample will have little regulatory consequence. If the adjusted and unadjusted portions agree (both are toxic or nontoxic), then the unadjusted alone would have had the same outcome as parallel testing. If the adjusted is toxic and the unadjusted is nontoxic, the unadjusted will be considered the most reliable because the acid or base will be assumed to have created artifactual toxicity not occurring in the receiving water. If the adjusted is nontoxic and the unadjusted is toxic, then there is a good indication of a pH effect or pH influenced toxicity, but this information, even though useful in a TI/RE, would not alter the determination based on the unadjusted sample that the effluent was toxic.

The purpose of whole effluent toxicity testing is to simulate the conditions which occur as the discharge enters the environment. These conditions include a gradient of both toxicant concentrations and pH as the discharge mixes with receiving water. The use of receiving water as dilution water mimics these conditions best. If the receiving water is nontoxic and free of diseases and parasites, then it may be used unless the permit specifies laboratory water.

If a lab believes that apparent effluent toxicity might be an artifact of a difference in pH between the test solutions and the receiving water, then the permittee may submit a request to switch to using ambient water as dilution water in future tests. Using ambient water as dilution water will produce pH conditions that are as close to the actual discharge situation as can reasonably be expected in a laboratory. If valid tests cannot be produced using ambient water as dilution water, then a request may be submitted to adjust the pH to match the pH at the edge of the mixing zone during critical conditions.

Control of pH rise in test solutions may be accomplished by holding test chambers in a CO_2 atmosphere or aerating with CO_2 (See *Environmental Toxicology and Chemistry*, Vol. 11, pp. 609-614, 1992). An oxygen headspace may be used to maintain adequate dissolved

oxygen levels without encouraging pH rise. More frequent test solution renewals may also be used to control pH drift. Addition of acid may not be used to control pH rise.

6. Randomization

A critical assumption in the statistical analysis of toxicity data by hypothesis testing is independence among observations. Independence of observations is especially critical for the parametric hypothesis test procedures (Dunnett's, Bonferroni's, and Student's t-tests) that are used for regulatory determinations. Randomization of test chambers is the method provided in all of the EPA test manuals for achieving independence of observations. Randomization of test chambers must be standard practice for labs conducting toxicity tests for NPDES permittees in this state. Randomization must be documented in the standard operating procedure (SOP) approved as a part of accrediting the lab for the test method. True randomization must be employed involving the use of random numbers to assign test container positions. The randomized bench sheets (hand written entries unless the balance automatically enters weights) must be submitted for all tests involving hypothesis testing. Failure to do so will cause test results to be rejected. (See Appendix A of any EPA chronic toxicity test manual or section 11.1.6 of the EPA acute manual.)

7. Tests Which Fail the Power Standards

Sometimes variability across replicates will prevent a large difference in response (in other words, a toxic effluent) from being detected as statistically significant. False negatives can happen when the number of replicates is low. The WET rule handles false negatives through the establishment of power standards. The WET rule contains both an acute statistical power standard and a chronic statistical power standard.

The acute statistical power standard says that acute toxicity tests must be able to detect a minimum of a 30 percent difference in survival between the ACEC and a control as statistically significant. The chronic statistical power standard says that chronic toxicity tests must be able to detect a minimum of a 40 percent difference in response between the ACEC or CCEC (the NOEC if the ACEC is unknown) and a control as statistically significant.

If a WET test does not meet the appropriate statistical power standard, then the permittee will be required to immediately resample the effluent and repeat the toxicity test with the number of replicates increased in order to meet the statistical power standard. (See Appendix E for an example calculation of compliance with the power standards.)

C. Other Testing Requirements

1. Dechlorination

WET tests conducted on effluent samples which are dechlorinated under any circumstance other than that allowed by WAC 173-205-080(3) or by the NPDES permit cannot be used

for regulatory determinations and must be repeated. We now prefer that samples for WET testing of chlorinated effluents be taken prior to the chlorinator if the ACEC is below 25% effluent and the discharge can meet water quality-based effluent limits for chlorine. Otherwise, WET testing must be performed on an unmodified sample of final effluent. See Appendix G, Chlorine Toxicity, for more explanation.

2. Acute Toxicity Test Duration

WAC 173-205-050(1)(c) requires that the duration of an acute toxicity test be 48 hours for an invertebrate and 96 hours for a fish. New permits will specify these durations for acute tests. Some older permits did not specify a duration for acute tests. When the permit has not specified acute test duration, then WAC 173-205-050(1)(c) should be followed or the toxicity test results might be rejected.

If an older permit specifies an acute test duration that is different than the durations in WAC 173-205-050(1)(c), the permittee should request that Ecology approve a change to the appropriate test duration. Acute test durations that are shorter than the durations in WAC 173-205-050(1)(c) could cause Ecology to require the permittee to repeat the effluent characterization for acute toxicity. Acute test durations, that are longer than the WET rule requires, penalize permittees unnecessarily.

3. Outdated EPA Manuals

Only the most recent version of an EPA manual should be used. For acute testing, it is EPA/600/4-90/027F. For freshwater chronic testing, it is EPA/600/4-91/002. For saltwater chronic testing with East Coast organisms, it is EPA/600/4-91/003. For saltwater chronic testing with West Coast organisms, it is EPA/600/R-95/136. All accredited labs were notified that tests initiated after April 15, 1996, must be conducted in accordance with these new manuals in order to be acceptable for effluent monitoring. These manuals can be obtained by calling the National Center for Environmental Publications and Information (NCEPI) at 513-891-6561 or downloaded from the Internet at ftp.epa.gov or gopher.epa.gov.

4. Reference Toxicant Tests

Reference toxicant testing must accomplish two purposes in the effluent monitoring program. One purpose is to evaluate test organism sensitivity, and the other purpose is to track lab performance of the test. Both purposes are best accomplished by a concurrent reference toxicant test conducted along with each batch of samples tested at the same time in a lab. Concurrent reference toxicant testing is the only method that produces a true positive control for a toxicity test. Concurrent reference toxicant testing with all tests is not required in the EPA manuals, but does represent a noteworthy commitment to quality assurance by any laboratory choosing to do so.

The minimum reference toxicant testing needed to meet our interpretation of the requirements in the EPA manuals (both sections 4.7 and 4.16) is one per month for every acute and 7-day (short-term) chronic test species used routinely (more than once per

month). Because an acute test result can be determined during a 7-day chronic test, acute and chronic reference toxicant testing for a fish or mysid can be combined. If a lab has difficulty establishing a concentration series that produces good results for both a lethal and sublethal endpoint, the lab may focus on lethality as long as the sublethal endpoint is not completely abandoned in the conduct and analysis of the test.

In addition to the nonroutine tests (test performed once per month or less), all tests conducted with bivalves, echinoderms, or plants are required to have concurrent reference toxicant testing. We require concurrent reference toxicant testing with each batch of samples tested with the bivalve development test, the echinoderm fertilization test, or the echinoderm development test. A group of tests qualifies as a batch if they are tested at the same time using gametes from the same spawning. Otherwise, additional concurrent reference toxicant tests are required. The bivalve and echinoderm tests are highly sensitive to the toxicity of many effluents. Lab technique is crucial. In addition, brood stock can vary in condition, and the concurrent check on test organism sensitivity is a good precaution. Spawnings are usually generous enough to supply concurrent reference toxicant tests. These tests often do not qualify as routine tests (more than once per month) anyway and would be required by the EPA manual to have a concurrent reference toxicant test. Algal toxicity tests must have concurrent reference toxicant tests for similar reasons. Concurrent reference toxicant testing is also required when test organisms (or the brood stock used to produce the test organisms) have been collected from the wild.

Section 4.7 contradicts itself somewhat on the frequency (monthly or concurrent with each test) of reference toxicant testing required when an outside supplier is used for test organisms. In choosing to require monthly as opposed to concurrent reference toxicant testing for routine (more than once/month) acute and 7-day chronic tests even when an outside organism supplier is used, we considered the following facts: The cultures of today's test organism suppliers are usually maintained at least as well as lab in-house cultures, and labs relying on in-house cultures are only required by section 4.7 to conduct a monthly reference toxicant test for tests conducted routinely. The routine test organism known to vary the most in control performance is *Ceriodaphnia dubia* and it is invariably cultured in-house by testing labs. Requiring concurrent reference toxicant testing for tests conducted routinely seems excessive when failures to meet control acceptability criteria cause many more routine test rejections than reference toxicant testing could. Increases in test costs, especially the cost of 7-day chronic tests, are to be avoided if possible. The alternative to concurrent reference toxicant testing in section 4.7 for labs getting test organisms from an outside supplier is reference toxicant testing by the organism supplier, and this alternative seems to be generally believed by testing labs as well as the Department of Ecology to be inferior to monthly reference toxicant testing by the testing lab.

Section 4.7 of the EPA manuals allows labs to evaluate the sensitivity of a batch of test organisms received from an outside supplier either by conducting concurrent reference toxicant tests with each acute or chronic test performed with effluents or by submitting reference toxicant data (control chart of at least five monthly tests) from organism suppliers. However, reference toxicant tests conducted by the supplier do not really provide reference toxicant test results that can be related to samples tested by the lab ordering the test organisms. In addition to the fact that organisms tested with reference toxicants by suppliers have not been packaged and shipped prior to testing, dilution water and other test

conditions are bound to differ between the supplier and the effluent testing lab. For these reasons, we do not accept the use by labs of reference toxicant tests performed by organism suppliers, and apparently labs agree because the vast majority have, to their credit, continued to conduct their own reference toxicant testing. Labs, however, should use organism suppliers who routinely conduct reference toxicant testing and control charting because, as noted in the table below, this information can be useful when deciding the consequences of lab conducted reference toxicant testing.

Section 4.16 of the EPA manuals (section 4.15 in the acute manual) requires labs to track the performance of every test method commonly done in the lab by conducting a monthly reference toxicant test that has the same test conditions (duration, endpoints, dilution water, etc.) as the effluent tests. We interpret "commonly" to mean the same as "routinely" in discussions of section 4.7 - testing performed more than once per month. If reference toxicant testing to evaluate the condition of test organisms required in section 4.7 of the EPA manual is performed as described above, then no additional reference toxicant testing need be done to evaluate ongoing lab performance of the tests. Control charting can be done with any appropriate reference toxicant test that was conducted to meet the requirements of section 4.7.

All labs must conduct ongoing control charting based on reference toxicant testing and report the results, acceptable or unacceptable, of the control charting in the report for each effluent or ambient water test. Acceptability is based on the standard test acceptability criteria for the test and on control charting with the upper and lower control limits set at twice the standard deviation (95 percent confidence) of the point estimates (LC_{50}, EC_{50}, IC_{25}, etc.) accumulated from the last 20 reference toxicant tests. At least five reference toxicant tests are needed to establish a minimally effective control chart for new tests. The reference toxicant test data must be presented with the report for each associated test.

Any reference toxicant test determined to be unacceptable must be repeated either until an acceptable result is obtained or until there have been three consecutive unacceptable test results (the initial unacceptable test plus two repeats). Because about 1/20 reference toxicant test results will fall outside of control limits due to chance alone, it is necessary to repeat unacceptable reference toxicant tests in order to reduce the role of chance. Assuming no unusual problems with test organisms or lab performance, there is only a 1/400 chance of two unacceptable reference toxicant test results in a row and only a 1/8,000 chance of three unacceptable results in a row. If a lab has no unusual problems, repeating an unacceptable reference toxicant test should quickly produce an acceptable result. If a lab repeatedly produces unacceptable reference toxicant test results, it will give confidence to the conclusion that the lab has problems with test organisms or testing technique. The EPA manuals ask that the frequency of occurrence be considered in the evaluation of unacceptable reference toxicant test results, and making this consideration when evaluating an unacceptable reference toxicant test will require the results of follow-up testing to determine the frequency of occurrence.

When the reference toxicant test result is within the 95 percent confidence limits, then the test report must state this fact and present the reference toxicant data at the end of the report. When the reference toxicant test result is outside the 95 percent confidence limits, then the test report must state this fact and present the reference toxicant data at the end of

the report. The lab should not delay test reports while waiting for the results of reference toxicant test repeats. The results from the first repeated test might be available in time for inclusion in the test report. If begun promptly, the results of all of the reference toxicant testing in response to an unacceptable reference toxicant test result will be available in time for the review of the test report. The WET Coordinator will contact the lab during the test review for any additional reference toxicant test data not contained in the test report.

When a reference toxicant test result falls outside of the 95 percent confidence limits, a lab must qualify the associated test result for an effluent or ambient water sample by a statement in the test report that the reference toxicant test result was outside control limits. The Department of Ecology WET Coordinator will decide whether these tests are acceptable based on the degree of departure from control limits and the frequency of occurrence. Because it is expected that an average of one out of 20 tests will fall outside of the control limits due to chance alone, the degree of departure from the control limits and frequency of occurrence will be considered before rejecting toxicity tests. Because control limits narrow as laboratory performance improves, the width of the control limits will also be considered before rejecting toxicity test results when the associated reference toxicant test results are just outside the limits.

The Biomonitoring Science Advisory Board (BSAB) criteria for acceptable intralaboratory variability provide values that are useful for considering the width of control limits while deciding whether to reject toxicity tests on the basis of reference toxicant test results. If the coefficient of variation (standard deviation ÷ mean toxicity value) from the reference toxicant test data used in control charting falls into the excellent (< 0.35) or good (0.35 to 0.60) range established by the BSAB, then a higher confidence in the test results is justified. If the reference toxicant test data coefficient of variation for the lab falls into the acceptable range (0.61 to 0.85), then a smaller amount of confidence should be applied. If the reference toxicant test data coefficient of variation for the lab falls into the unacceptable range (> 0.85), then none of the lab's test results are acceptable. Labs must report the coefficient of variation for the last 20 reference toxicant tests in every report for the same test conducted on an effluent or environmental sample.

Reference:

Biomonitoring Science Advisory Board. BSAB Report #1, ***Criteria for Acceptable Variability of Marine Chronic Toxicity Test Methods.*** Washington Dept. of Ecology. February 1994.

Effluent or ambient water toxicity test results will be accepted or rejected based on the following table. Rejection will occur when any condition in the appropriate "Test Accepted" box was not met or when any condition in the appropriate "Test Rejected" box was met.

Table for Determining Test Rejection Based on Reference Toxicant Test Results		
Unacceptable Reftox Tests	Test Accepted	Test Rejected
Only the original reftox test result was outside of control limits (the first repeat reftox test result fell within control limits)	If the organism supplier reftox results were within control limits, and the coefficient of variation for the last 20 reftox tests is ≤ 0.85	If there are notable reporting errors or deviations from test protocol, or if the reftox test result fell outside of control limits to the more sensitive side (point estimate was too low) by 3 or more standard deviations and the effluent test showed toxicity at levels of regulatory concern
Both the original and the first repeat reftox test results were outside of control limits (the second repeat reftox test result fell within control limits)	If the 95% confidence interval for the point estimate used in control charting can be calculated and in both failing reftox tests overlapped the control limits in the control chart, organism supplier reftox results were within control limits, and the coefficient of variation for the last 20 reftox tests is ≤ 0.60	If there are notable reporting errors or deviations from test protocol, or if any reftox test result fell outside of control limits to the more sensitive side (point estimate was too low) and the effluent test showed toxicity at levels of regulatory concern
All three reftox tests were outside of control limits	never	always
Coefficient of variation for the last 20 reftox tests > 0.85	never	always

Effluent tests and their associated (initial) reference toxicant tests must have start dates separated in time by no more than 18 days. It makes no sense to use a monthly reference toxicant test to evaluate lab performance for the next 30 days when tests conducted the previous week are closer in time than those toward the end of the 30-day period. Labs typically take about two weeks to produce a test report. From the point of view of practicality and the most meaningful control charting, it makes sense for a reference toxicant test result to be used retroactively about two weeks. The reference toxicant test result will then be used for control charting for the balance of the monthly time period. A grace period of 7 days will be added to the 18 days for tests begun from December 1st to the following January 10th. Acute tests will be allowed a grace period of 4 days over the 18 day maximum.

Because point estimates provide the best basis for control charting, all labs must control chart using point estimates. Point estimates require fewer replicates than NOECs and reference toxicant testing may be done using the minimum number of replicates allowed by the test method.

Another Ecology staff person with primary responsibility for reference toxicant testing requirements is the Advisory Laboratorian in the Quality Assurance Section who reviews standard operating procedures (SOPs) for toxicity tests and accredits labs. For bioassay labs to maintain Department of Ecology laboratory accreditation, the QA section has begun to require participation in a round-robin test (such as the DMR-QA) or the performance of one reference toxicant test at least once every six months. Except for bioaccumulation/bioconcentration tests such as ASTM E 1022 and E1706, this requirement applies to all effluent, sediment, soil and dangerous waste characterization type bioassay methods for which labs seek continued accreditation. In the event that a lab does not conduct any tests on environmental samples using a particular species/method within a six-month period, it must perform a reference toxicant or round-robin test. In the event that a lab does not conduct any tests by a particular method within a one-year period, it must do two reference toxicant or round-robin tests for that year. Further, these tests must be done at least four months apart. This is to assure that the labs maintain proficiency with the species and methods for which they are accredited. The Quality Assurance Section can efficiently enforce good reference toxicant testing requirements because they have direct authority over labs, approve SOPs, and conduct routine onsite audits. The Water Quality Program will also consider QA Section approval in our assessment of reference toxicant testing requirements. The telephone number of the QA section is 360-895-4649.

5. Outliers

Labs may identify outliers if they choose to do so using an appropriate statistical procedure (Gentleman Wilk's A statistic, Dixon's test, etc.) and submit the tests results with the outliers both excluded and included. If outliers are to be excluded, then they should be identified at both low and high ends of test organism performance. An important function of the WET database is to provide an accurate record of test performance as well as effluent toxicity, and the exclusion of outliers will hide some important features of test performance. Most labs are likely to continue to not look for outliers and include the results from all test chambers in the calculations, and this is also how we will be recording most test results. However, outlier identification is considered useful in the following three circumstances:

- The lab has a physical explanation (fish accidentally siphoned but not killed outright, contaminated glassware, temperature excursion, etc.) for one or two aberrant values and wishes to officially exclude the results from those test chambers. Test organisms which were accidentally killed by a documented physical event do not need to be identified as an outlier in order for the start count to be reduced (single mortalities) or the replicate to be dropped from calculations (complete loss of a test chamber). Outlier identification is not a solution for sporadic mortalities as discussed below in section 8. Sporadic Mortalities.

- If the lab and permittee choose to do so, outlier identification may be used to meet the power (statistical sensitivity) standards when the pooled variance has been adversely affected by one or two values. Otherwise, outlier identification should not be used to suppress test variability and bias hypothesis testing.

- If the lab and permittee choose to do so, outlier identification may be attempted to improve the concentration-response relationship of a test rejected for being anomalous.

If outlier identification provides an acceptable concentration-response, then the test need not be repeated.

6. Excessive Time to Produce a Test Report

The WET Rule contains time limits for permittees to respond to different circumstances involving toxicity test results. Labs should be careful not to take more than four weeks after completing a test to produce the test report or risk adding to permittee difficulties. Timely test reports are especially important as WET limits become common. Labs should give the permittee an immediate telephone call if serious toxicity has occurred and the test report is a month away. We will continue to track the time it takes labs to produce a report and may eventually produce a comparative table of lab turn-around times.

7. Aeration of Test Chambers

In addition to being kept to the minimum duration necessary to maintain desired dissolved oxygen levels, aeration in test containers after test initiation must not be initiated more than once if it can be avoided. Aeration in test containers should be continued long enough for dissolved oxygen to remain above the minimum level until test solution renewal or test termination. As a measure to avoid having to repeatedly initiate aeration of test chambers, the sample should be aerated a little longer prior to test solution renewal if maintaining dissolved oxygen levels has been a problem during the test.

Use of an oxygen headspace would be preferable to aeration in maintaining adequate dissolved oxygen because it is nonintrusive to the test solutions.

8. Sporadic Mortalities

Sporadic mortalities are deaths of test organisms that are not related to sample toxicity and do not fit a good concentration-response relationship. These sporadic mortalities sometimes cause a flat concentration-response relationship with nearly equal proportions alive which resemble an infection rate not toxicity. At other times, sporadic mortalities are confined to a few test chambers scattered throughout the test as if susceptible individual test organisms were becoming infected and concentrating the pathogen within their test chambers causing large standard deviations in proportion alive in those concentrations. Inadequate cleaning or rinsing of glassware and poor quality disposable test cups can also cause sporadic mortalities. Regardless of cause, anomalous test identification criteria 2 and 5 identify the occurrence of these sporadic mortalities and provide labs with an opportunity and incentive to improve test performance. Sporadic mortalities are a common and preventable cause of anomalous test results.

If sporadic mortalities have been occurring, then a lab should give extra attention to proper glassware cleaning and rinsing so that toxic residues are removed. Using only food grade disposable cups and changing supplier when there is a problem can reduce sporadic mortalities. Labs should not skip steps in the test method which involve quality control of test chambers such as those which call for soaking test containers in water overnight prior to test initiation. Running acute tests with fathead minnows or daphnids at 20° C instead of

25° C might reduce the occurrence of sporadic mortalities. Keeping samples at 4° C from the moment of collection until used in the test might also reduce mortalities due to pathogens.

Pathogens which will infect test organisms can come from inside a lab, from a composite sampler, or from the sample itself. These pathogens can often be observed as filaments or patches on test organisms. An alert lab will notice whether diseases are killing test organisms and look for a source. If sporadic mortalities tend to occur mostly with a few clients, then the source of pathogens is likely the effluent or composite sampler. If sporadic mortalities occur for all clients, in controls, or in reference toxicant tests, then the source of pathogens is likely within the lab.

Cleaning, rinsing, and disinfection should be thorough and routine for all reusable glassware, all organism holding containers, and all general lab surfaces such as bench tops and the insides of refrigerators and incubators. Test chambers should be kept covered to prevent airborne transfer of microbes. Adult mosquitoes, chironomids, and other flies must not be allowed free in the lab. Enough sterile pipettes or other equipment for transferring test organisms from chamber to chamber should be used so that cross contamination between replicates does not occur.

Composite samplers and their tubing make ideal surfaces for growing microbes which might infect test organisms. Composite samplers should have all tubes changed and be cleaned before sampling for toxicity testing.

The EPA manuals recommend that unhatched *Artemia* cysts and empty exoskeletons not be fed to fathead minnow larvae. Regular and thorough cleaning and disinfection of Artemia hatcheries can eliminate pathogens which might cause sporadic mortalities.

Some effluents are associated with sporadic mortalities more often than others. Noncontact cooling water has the highest frequency of sporadic mortalities. Ambient samples can also have sporadic mortalities. Naturally occurring pathogens are likely the cause of sporadic mortalities in ambient water. Pathogens in noncontact cooling water might originate in the natural water source for the cooling water and sometimes be enhanced by growing in pipes or other surfaces within the plant. *Environmental Toxicology and Chemistry* has published two informative articles on pathogens in toxicity tests; one in Vol. 15, No. 5, pp. 761-764 and the other in Vol. 16, No. 2, pp. 351-356.

If an effluent from a permittee regularly produces sporadic mortalities, a lab may ask for permission to ultraviolet (UV) disinfect that permittee's samples. If our database shows regular sporadic mortalities for the permittee and shows that the lab does not have a general problem with sporadic mortalities, then UV disinfection will be allowed. Copies of the permittee's records of composite sampler maintenance must be submitted with the request to conduct UV disinfection of samples. Little is known at this point about the UV exposure necessary to eliminate sporadic mortalities caused by pathogens except that it should be kept to the minimum necessary and that the minimum exposure necessary is less than that reported in the papers mentioned in the preceding paragraph. One lab here in the Northwest has been routinely disinfecting noncontact cooling water and river water successfully using an UV exposure of about 2 minutes (1 or 2 passes through the unit

water in order to minimize any concentration gradient of the artificial salt in the test concentrations. The State of Washington prefers that each test be performed with a single source of salinity either artificial salts, hypersaline brine, or a combination of a natural seawater and a brine prepared from the same natural seawater.

11. Deviations from Protocols and Acceptability Criteria

Deviations from the protocols or failures to meet control acceptance criteria need not always cause test rejection. As a reward for honesty and accuracy, tests will be occasionally accepted even if the protocol was not completely followed or if the control did not meet performance criteria. The test results must indicate no significant toxicity. Protocol deviations must be both minor and not likely to mask toxicity such as small temperature excursions or the use of the wrong size test chamber. Control acceptability criteria failures must be accompanied by robust and consistent organism performance at all other test concentrations.

In order to have an imperfect test result accepted, a lab must call Randall Marshall at 360-407-6445 either during or immediately following the test. After telephone permission has been given, the lab must completely document the test conditions and the telephone conversation in the test report. If the lab makes few requests and has demonstrated a willingness in the past to repeat imperfect tests, the permission may be granted and the test report accepted.

12. Dual Endpoint Tests

Labs sometimes provide their clients with acute test results from a 7-day chronic test. This is sometimes called "dual endpoint testing." To have a dual endpoint test, the daily survival from a 7-day chronic test at 48 hours (daphnids or mysids) or 96 hours (fish) is used as the final count in an acute test. Permittees should always be informed by the lab when dual endpoint testing will deprive them of the advantages of a separate acute test run at a cooler temperature, without daily renewals, or using older and larger test organisms.

Acute tests derived from 7-day *Ceriodaphnia* chronic tests are not acceptable because this chronic test fails to meet the minimum number of test organisms required per test chamber and per test concentration for an acute test. Fisher's Exact Test is also not acceptable for analyzing the results of an acute test and the setup (one organism per test chamber) of the *Ceriodaphnia* chronic test makes Fisher's Exact Test the only option.

D. Check for Completeness of Report

1. Paper Submittals

Labs must attach a readable copy of all bench sheets and chain-of-custody forms to the WET test report. The bench sheets must include both the toxicological and water chemistry data for both the WET test and reference toxicant test. The bench sheets must contain actual counts (not percentages) in order to be acceptable. Start counts must be clearly

recorded on the bench sheet. The WET test report must include computer printouts of test data and statistical analyses.

The test report must contain all of the information needed for comparison with the requirements below in Chapter III. Toxicity Test Report Checklists. The sample date (ending date for composite samples) and sampling method (grab or composite, volume, sample container size and material, temperature of sample, etc.) must be reported somewhere in the test report or chain-of-custody form. Test organism source, age, and unusual conditions (lethargy, hyperactivity, spots or filaments, discoloration, excessive ventilation, etc.) must be reported. The report must contain a description and justification of any dechlorination procedure used. The stoichiometric calculations for determining the proper amount of dechlorinating agent must be included in the test report. The report must contain a description and justification of any sample filtration procedure used. The report must contain a description and justification of any aeration or pH control/modification used during the test. Any special circumstances such as treatment system upsets known to exist at the time of the sample must be reported. Each test report must contain a section for noting deviations from test protocol where all deviations must be accurately listed or the absence of deviations noted.

The test report will be reviewed for inconsistencies and typographical errors. Examples of report inconsistencies include referring to different test species (or different test methods) on different pages of the report. Examples of typographical errors include data entry errors or transposing the sample date and test date. Labs will be contacted directly about occasional report inconsistencies or typographical errors. If these inaccuracies occur more often than occasionally, then permittees will be contacted to resolve the problem.

2. Electronic Submission of Test Data

The Department of Ecology will be making the submission of WET test results and reference toxicant test results on computer floppy disks (3.5" is best) voluntary. New permits will instruct permittees to forward any floppy disks provided voluntarily by the lab. Those existing permits which contain a requirement for electronic submission will not change, and permittees must meet this requirement.

If an efficient system can be worked out between a lab and the WET Coordinator, monthly reference toxicant test results may be submitted once electronically for all of the associated tests on effluents or ambient water in that month. The individual test reports can then be smaller and still be considered complete. The individual test reports must still summarize the results of the reference toxicant test and control charting but need not contain the reference toxicant test data. The test number for the reference toxicant test must be clearly identified in the report for every associated test. Concurrent reference toxicant test results may also be submitted electronically. Be aware that only about half of electronic submissions are working right now so achieving an efficient system may take some effort.

The codes for electronic submission to the Department of Ecology database are inconsistently used. All in-house cultures in all states are identified as XXIH. Hatching or spawning organisms in the lab do not constitute in-house culture if the eggs or adults were

obtained from outside the lab. For the organism source code used in data entry and electronic submission, the source of the fish is considered to be the facility which maintains the brood stock and produces the fertilized eggs. Static tests are defined as tests with no renewals. Static-renewal tests are tests with one or more renewals. Use the code CA0000000 if you do not know your client's permit number. The test material codes for stormwater are SRW1 (municipal) and SRW2 (industrial). Some industries have test material codes specific to their effluent such as pulp mills (EFF5), oil refineries (EFF6) and aluminum smelters (EFF7). The new EPA manuals are coded: EPAA 91 (acute manual), EPAF 94 (freshwater chronic manual), EPAM 94 (East Coast marine manual), and EPAW 95 (West Coast marine manual).

III. Toxicity Test Report Checklists

A. Sample Handling

1. Transfer and Storage

Sample transfer must be documented with signed and dated chain-of-custody forms which must accompany the test report. For composite samples, the sample date is considered to be the end date of the compositing period. As described in the next section, 2. Sample Holding Time and Temperature, samples must be immediately chilled usually to 4°C. Composite samples are chilled as collected and grabs immediately following collection. Labs must store samples at 4°C in the dark with no headspace.

Labs which go to the extra effort and expense to use glass containers provide superior sample protection and preservation. Minimization of head space is also important with glass containers. All glass containers should be filled to the top with sample. A sample should be collected into two or three glass containers of an adequate size for daily renewal. These must be stored at 4° C in the dark.

Chain-of-custody forms must accompany all samples unless:

1. A person from the testing lab does the actual sampling and then delivers the sample personally to secure storage at the lab; or

2. Personnel who are all employees of the organization which is also the discharge permit holder are the only ones conducting the sampling, transportation and toxicity testing, and a responsible person from that organization signs a page in the test report stating that the result is a honest and accurate reflection of the toxicity of the sample.

Chain-of-custody forms must contain the name and address of the discharger, date and time that the sample is taken (beginning and end if a composite), the name of the sampler, the type of sample (outfall #, grab or composite, effluent or stormwater, etc.), and the number and volume of sample containers. The chain-of-custody form must describe the type of sample container.

The sampler's signature must be in the first "relinquished by" blank. Each person subsequently taking physical custody of the sample must sign the next "received by" blank and then the next "relinquished by" blank when the sample is given to someone else. This sequence of signing is repeated until the sample is secure at the testing lab. Every signature must have a date and time, and each pair of "relinquished by" and "received by" signatures must have the same date and time (within a couple of minutes to allow for differences in watches or clocks). The use of a courier is the only circumstance when a pair of "relinquished by" and "received by" signatures can have significantly different times.

Couriers do not need to sign the "received by" blank on the chain-of-custody form if the cooler containing the samples was packed by the sampler and has been locked or sealed with a seal that is initialed and dated by the sampler and cannot be removed without the removal being obvious (i.e. evidence tape). The name of the courier company and the method for locking or sealing the cooler must be identified on the chain-of-custody form. The sampler signs and dates (including time) the "relinquished by" blank on the chain-of-custody form and immediately locks or seals it in the cooler with the samples. Immediately upon receipt at the testing lab, a responsible person inspects the cooler to make sure that locks or seals are intact, opens the cooler, removes the chain-of-custody form, signs and dates (including time) it, and places the sample containers in secure storage at 4° C (unless the test is begun immediately). When a courier is used, all signers to the chain-of-custody form are testifying to the proper condition of the cooler, lock, or seal unless otherwise noted on the form.

One chain-of-custody form should accompany the sample throughout its travels. When a second lab is subcontracted to perform some of the tests on a sample originally received at the primary testing lab, the required chain-of-custody procedure is:

1. The sampler completes all information pertinent to sampling and transportation on the chain-of-custody form and signs relinquishing the sample. The chain-of-custody form is locked or otherwise sealed in the cooler if a courier is used.

2. The primary lab opens the cooler immediately upon receipt, signs the "received by" line on the chain-of-custody form, and makes a copy for inclusion with their test report.

3. The primary lab notes on the original chain-of-custody form the number and volume of containers placed in the cooler for the second lab, notes the method of transportation, signs the second (or next) "relinquished by" line, and locks/seals the form in the cooler with the sample.

4. The second lab opens the cooler immediately upon receipt, signs the next "received by" line on the chain-of-custody form, and makes a copy for inclusion with their test report.

5. The completed original chain-of-custody form is returned to the primary lab to be kept in their records.

2. Holding Time

Maximum holding time from sample collection to test initiation is 36 hours. In order to be able to see if the holding time is exceeded, the date and time of test initiation must be clearly recorded on the bench sheet and a copy included in the test report.

If the sample is received at the testing lab within one hour after collection, is a grab sample, and is immediately refrigerated at the lab or used in a test, it must have a temperature between 4° C and 20° C. If the sample is received at the testing lab within 4 hours after collection, it must be between 4° C and 12° C. All other samples must be between 4° C and 8° C. Sample temperature must be measured by the lab at receipt and recorded on the chain-of-custody form or initial water chemistry form. Samples must be stored at 4° C until used in a test. Tests conducted on samples received too warm will be rejected. Tests

conducted on samples received much below 4° C will be accepted, but the test review will contain warnings about the consequences of frozen samples. If a 7-day chronic test is already underway using an initial sample that was a good temperature at receipt and the second or third samples arrive too warm, use the process in section II. C. 11. Deviations from Protocols and Acceptability Criteria to find out whether to continue the test or not.

Based on experience, setting a maximum temperature limit of 4° C for all samples is not practical and may not be necessary. However, 4° C is the best temperature for sample preservation and will remain the ideal temperature set as a goal for all samples. A typical sample is taken by an automatic composite sampler which should hold the sample at 4° C while being collected. The sample is then placed in a cooler with ice to keep the temperature low during transportation, but sample temperature often rises anyway especially if gel packs ("blue ice") are used. Real ice cools well, but must be used in moderation and be from a verifiably clean water source. It is not a good idea to have a cooler arrive with sample containers sloshing around every which way in slushy melt water. We need to set a maximum temperature that a 4° C sample can rise to and still be adequately preserved for 36 hours. 8° C is about the same temperature used in public health regulations as the maximum refrigeration temperature (45° F) to prevent food spoilage for a short period of time. Effluent samples should be reasonably preserved if kept at 8° C or below for a maximum of 36 hours. One lab already rejects samples received above 8° C, and we believe that all other labs must begin doing so in the interest of accurate and fair test results.

Samples quickly transported to a lab need less preservation than samples taking most of the 36-hour holding time in transit. 12° C is a relatively cool but common ambient air or water temperature in this state and a sample should not warm above this temperature during four hours of transportation. Composite samples should still be close to 4° C after four hours and we will suspect a defective composite sampler if samples arrive at the lab within four hours above 8° C and note this suspicion on the test review (but still accept the test if the sample was 12° C or less). Except on the hottest days, most grab samples should cool to 12°C while being transported for four hours with a generous amount of ice. On especially hot sunny days or if the effluent itself is warm, grab samples may need to be cooled in an ice bath onsite prior to packaging for shipment.

Grab samples arriving at a lab within one hour of collection will often not have the time to be cooled to 4° C. 20°C is a typical but not particularly warm room temperature, and therefore makes a good maximum temperature for a grab sample transported to a lab on ice within an hour of collection.

Freezing of samples during transportation does happen on rare occasions. Freezing usually pops the cap off of the sample container or bursts the container. Freezing will concentrate dissolved solids (as in making brine) and induce big changes in dissolved gases as a sample cools, freezes, and then thaws. Samples which have frozen must not be used for effluent toxicity characterization or compliance tests.

The original sample may be used for test solution renewal at 48 hours in an acute test if stored at 4°C in the dark with no headspace.

If a chronic test requiring daily renewal will be conducted on an intermittent discharge which does not allow the collection of three separate samples over seven days, then sufficient sample must be collected during all of the available discharge events to provide daily renewal. The extra sample must be collected in a separate container with no headspace. It must be stored at 4° C until used according to the schedule in the EPA test method.

3. Filtration

No filtration of samples is allowed unless the necessity for filtration has been documented. Justification for filtration should be based on the observation of organisms that would attack, be confused with test organisms, or otherwise interfere with the test. Most samples do not contain indigenous organisms that would attack or be confused with test organisms. Many labs rarely filter samples and have no problems with toxicity tests. Unless the test report contains good justification, a lab will have tests on filtered samples rejected.

If a lab can demonstrate that a particular effluent contains organisms which interfere with toxicity testing, then samples of that effluent may be filtered. A good demonstration would be to conduct a toxicity test with twice as many replicates at 100 percent effluent with half of the replicates filtered and half unfiltered. If there is a difference in test results and organisms are identified in the filter backwash, then filtration of that effluent has been justified. This demonstration need only be made once for each effluent discharge and then all future samples may be filtered. The demonstration is not required in order to filter samples of surface water or samples from treatment lagoons with retention times in excess of two days if the lagoon is part of a biological treatment system or has been colonized by aquatic plants.

Filter pore diameters should be no smaller than is necessary to remove the unwanted organisms. Pore diameters must never be smaller than specified in the test method (60 µm except for *Selenastrum* which is 0.45µm).

4. Aeration

No aeration of samples is allowed unless justified by measurements showing dissolved oxygen to be at concentrations considered deleterious. Dissolved oxygen measured at concentrations below 4.0 mg/L (6.0 mg/L for rainbow trout) justify aeration.

Supersaturation of dissolved gases in the sample would justify aeration only after preparation of test concentrations and pouring of the replicates have been shown to not remove or dilute excess gases adequately. The manipulation of test solutions alone can often remove or dilute supersaturation sufficiently. The replicates for the 100% effluent concentration should be prepared first so they can equilibrate while the effluent dilution series is prepared and the replicates poured. If this procedure occasionally does not work, then the test chambers should be aerated. If this procedure often fails to work for samples from a discharge, then document the problem and request permission to routinely aerate samples from the discharge prior to test setup.

In addition to being kept to the minimum duration necessary to maintain desired dissolved oxygen levels, aeration in test containers after test initiation must not be initiated more than once if it can be avoided. Aeration in test containers should be continued long enough for dissolved oxygen to remain above the minimum level until test solution renewal or test termination. Try aerating the sample a little longer prior to test solution renewal if maintaining dissolved oxygen levels has been a problem during the test. If extended aeration of the sample does not work, then aeration of all test chambers should begin and continue until test termination. When aerating test chambers, aerate all test chambers including controls and test chambers which have adequate dissolved oxygen levels.

B. Water Quality Measurements

1. Purpose

Water quality measurements are important mainly for labs to use in monitoring and controlling test conditions. The test methods require these measurements for this reason. These measurements can also aid in test interpretation, but the biological data are the major influence on the determination of test quality. The following parameters and schedule must be followed for all toxicity tests whether acute or chronic. The list also notes those circumstances where water quality measurements will affect test acceptability.

Echinoderm and bivalve tests are exceptions to the water quality measurement schedule below. All parameters are measured, but because test chambers are too small to allow the measurements, there are differences in the schedule. The water quality measurements for the echinoderm fertilization test must be done at test initiation in the test chamber stocking solutions. The water quality measurements for the bivalve and echinoderm development tests must be done at test initiation and termination in a single extra replicate vial that has been setup specifically for the water quality measurements at each concentration and the control and used for these measurements.

2. Parameters and Schedule

Temperature: Measured in at least five test chambers (one on each edge and one near the middle) at the beginning of a test, daily during the test (before renewal if solutions are renewed that day), and at test termination. Experience has shown that inadequate monitoring and maintenance of temperature contribute to poor control performance and to test variability. Temperature must be measured in test chambers or in surrogate test chambers distributed throughout the test chambers. Failure to adequately measure and control temperature will cause test results to be rejected. After a track record for temperature control has been established for a test measuring as described above, then a request may be made to reduce the requirement. *Ceriodaphnia* chronic tests conducted in water baths will not have the temperature monitoring requirement reduced.

Dissolved Oxygen: Dissolved oxygen should be measured in the control and in at least one test chamber at every effluent concentration once per day at a minimum and often enough to detect any drop in dissolved oxygen before test organisms are adversely affected. Dissolved oxygen must be measured in one test chamber at each effluent concentration at

test initiation in order to determine if aeration is necessary to achieve the desired dissolved oxygen concentrations (or remove supersaturation). Dissolved oxygen should be checked again several hours later to see if it has dropped sufficiently to cause concern. If it has dropped significantly, then dissolved oxygen should be measured more often than daily. If dissolved oxygen does not drop significantly, then it may be measured once per day after any test solution renewal for the day. Dissolved oxygen measurements are required in order to justify aeration of the sample or test chambers. Test results will be rejected if aerated is done when not justified or if dissolved oxygen is allowed to persist at levels lower than that specified in the test method.

pH: Measured in the control and in at least one test chamber at every effluent concentration at the beginning of a test, daily during the test (before renewal if solutions are renewed that day), and at test termination. In order to provide information on pH changes during sample storage prior to renewals, pH must also be measured, at a minimum, in 100% effluent after test solution renewal. pH differences between concentrations or over time should be noted in the test report.

Conductivity: Measured in the dilution water and 100% effluent at the beginning of a test using freshwater organisms, at test solution renewal, and at test termination.

Salinity: If the effluent has salinity nearly equal to the dilution water and no brine or artificial salts are used in a test involving saltwater organisms, salinity is measured in the dilution water and 100% effluent at the beginning of the test, at test solution renewal, and at test termination. Salinity is measured in the dilution water and in at least one test chamber at every effluent concentration at the beginning of a test using saltwater organisms, at test solution renewal, and at test termination. Test results will be rejected if the salinity is not maintained within accepted ranges equally in all test concentrations

Total Hardness: Measured at the beginning of a test using freshwater organisms in the dilution water and 100% effluent.

Total Alkalinity: Optional at lab discretion. Recommended, but no longer required.

Total Ammonia: Measured at the beginning of the test in all samples which might contain ammonia and at any test solution renewal using fresh sample (all municipal effluents and any industry with the potential for ammonia). Caution should be exercised so that permittees do not have to pay for a toxicity identification evaluation to discover that ammonia was the cause of noncompliance.

Total Residual Chlorine: Measured at the beginning of a test in all samples which might contain chlorine and at any test solution renewal using fresh sample (all municipal effluents and any industry with the potential for chlorine). Measured in the dilution water at the beginning of all tests and at test solution renewal in all tests where tap water is used. Caution should be exercised so that permittees do not have to pay for a toxicity identification evaluation to discover that chlorine was the cause of noncompliance.

C. Toxicity Tests and Species

1. Acute Toxicity Tests and Species

The WET rule requires that effluents with a risk for aquatic toxicity are tested at a minimum for toxicity to a fish, an invertebrate, and any appropriate plant. Because EPA has not provided any test for acute toxicity to plants, effluents can be tested for acute toxicity only with a fish and an invertebrate. Acute toxicity tests with fish are 96-hour static-renewal tests. Acute toxicity tests with invertebrates are 48-hour static tests. A lab may provide daily feedings, if necessary, in any acute toxicity test as long as each feeding is followed by an 80% test solution renewal using either a fresh effluent sample or one stored at 4°C. Labs have the option of very gently aerating daphnid test chambers if dissolved oxygen levels fall below the values in the following table.

Daphnids are the invertebrate species for acute toxicity testing. The fathead minnow (*Pimephales promelas*) is the recommended acute WET testing fish species for all permits. EPA has developed the freshwater WET testing program around the use of fathead minnows for fish testing. If Ecology decides to require acute WET testing with rainbow trout (*Oncorhynchus mykiss*) in order to provide direct protection of salmonids, it is likely that the permit will also require fathead minnow testing so that any TI/RE can be performed with fathead minnow. A correlation between the sensitivities of the two fish can be established during effluent characterization for use in guiding the TI/RE.

Because of occasional shortages of rainbow trout of the correct age for testing, we will begin accepting brook trout (*Salvelinus fontinalis*) as a substitute using the same test conditions listed in the table below for rainbow trout except that the age range is 30 to 60 days post-hatch and at least 2 days past swim-up. Be sure to check to get your client's OK before making the substitution.

If the effluent itself is freshwater, freshwater species will be used for acute WET testing regardless of the salinity of the receiving water. If the effluent is too saline for freshwater organisms, the permit will require acute testing with the silverside minnow (*Menidia beryllina*) and a mysid (*Mysidopsis bahia*). Topsmelt (*Atherinops affinis*) or the West Coast mysid (*Holmesimysis costata*) may be substituted as long as organism age, test solutions and containers, number of replicates, number of organisms/chamber, test temperatures, and salinity are in accordance with the tables below in part III.C.3. Standard Saltwater Chronic Toxicity Tests. The East Coast mysid and silverside salinity should also be in accordance with the tables in part III.C.3. below.

If salinity adjustment is needed, artificial sea salts must be used in acute toxicity testing because the WET rule requires that the response in 100 percent effluent be used to determine the need for an acute toxicity limit or a new effluent characterization.

All conditions in the table, Acute Toxicity Test Required Conditions, on the following page must be met and reported for each toxicity test.

Table of Required Acute Toxicity Test Conditions

test organism	test type	chamber size	solution volume	# organisms per chamber	# replicates	age	temperature
Ceriodaphnia dubia	48-hr static	minimum 30 mL	minimum 15 mL	minimum 5	minimum 4	< 24 hrs	20° ± 1°C or 25° ± 1°C
Daphnia pulex/magna	48-hr static	minimum 30 mL	minimum 25 mL	minimum 5	minimum 4	< 24 hrs	20° ± 1°C or 25° ± 1°C
Pimephales promelas	96-hr static-renewal (at 48 hrs)	minimum 250 mL	minimum 200 mL	minimum 10	minimum 2 (eff. char.) 4 (compliance)	1- 14 days, 24 hr range in age	20° ± 1°C or 25° ± 1°C
Oncorhynchus mykiss	96-hr static-renewal (at 48 hrs)	minimum 5 L	minimum 4 L	minimum 10	minimum 2 (eff. char.) 4 (compliance)	15 - 30 days after swim-up[1.]	12° ± 1°C
Menidia beryllina	96-hr static-renewal (at 48 hrs)	minimum 250 mL	minimum 200 mL	minimum 10	minimum 2 (eff. char.) 4 (compliance)	9 - 14 days, 24 hr range in age	20° ± 1°C or 25° ± 1°C
Mysidopsis bahia	48-hr static-renewal (at 24 hrs)	minimum 250 mL	minimum 200 mL	minimum 10	minimum 2 (eff. char.) 4 (compliance)	1 - 5 days, 24 hr range in age	20° ± 1°C or 25° ± 1°C

NOTE: All of these table items and general items must be documented in each test report.

1. See Appendix A for a complete discussion of trout age determination. If brook trout are tested, age is 30 - 60 days post-hatch a

2. *Menidia beryllina* may be fed daily as long as an 80% renewal of test solution follows 2 hours after each feeding.

GENERAL ITEMS

The only approved test manual is EPA/600/4-90/027F.

Illumination must be for 16 hours at 10 - 20 µE/m^2/s (50 - 100 ft-c) followed by 8 hours of darkness.

Holding time is 36 hours maximum prior to test initiation. Renewals may be made using the original sample after 36 hours as long a

Controls must have at least 90% survival or the test should be repeated as soon as possible on a fresh sample.

2. Freshwater Chronic Toxicity Tests

Chronic WET test selection is fairly simple for discharges to freshwater. EPA recommends testing with a fish, an invertebrate, and a plant and has provided only one of each for freshwater chronic WET testing (fathead minnow, *Ceriodaphnia dubia*, and *Selenastrum capricornutum*). WAC 173-205-050(1)(a) requires that effluents with a risk for aquatic toxicity be tested at a minimum for toxicity to a fish, an invertebrate, and if appropriate, a plant. Permits for discharges to freshwater will contain standard requirements for the use of fathead minnow and *Ceriodaphnia* in chronic toxicity tests. The fathead minnow chronic test will measure survival and growth. The *Ceriodaphnia* chronic test will measure survival and reproduction.

Selenastrum is considered a supplemental chronic toxicity test. *Selenastrum* is often less sensitive than fish and invertebrates in WET tests. In addition, *Selenastrum* tests suffer from various effects which can mask or confuse the measurement of effluent toxicity. However, any clearly toxic response in an effluent test using *Selenastrum* is a good indication of toxicity to plants, and it will sometimes be required.

All conditions in the following tables for the freshwater chronic toxicity tests must be met and reported for each test. The standard chronic tests require three separate samples for renewals in a 7-day chronic test.

Ceriodaphnia Survival and Reproduction

Test species: *Ceriodaphnia dubia*

Approved test method: EPA/600/4-91/002

Test type: 7-day static-renewal (> 90% renewal of test solution in each test chamber daily by transfer of test organism to another container with fresh test solution)

Temperature: 25° ± 1°C

Illumination: Illumination must be for 16 hours at 10 - 20 µE/m^2/s (50 - 100 ft-c) followed by 8 hours of darkness.

Test chamber size: 30 mL (minimum)

Test solution volume: 15 mL (minimum)

Age of test organisms: < 24 hours and within an 8 hour age range

Number of organisms/chamber: 1

Number of replicates/concentration: 10 (minimum)

Feeding: 0.1 mL YCT and 0.1 mL algal suspension daily

Aeration: none unless DO < 2.0 mg/L and then is optional at lab discretion using a very low bubbling rate

Test duration: The duration of exposure is expressed in terms of time (seven days) for the survival endpoint and in terms of life cycle (three broods) for the reproduction endpoint. Final survival counts must be taken at the end of 7 days. Final counts of neonate production should be taken immediately upon production of the third brood by 60% of the surviving control organisms. The third brood will commonly occur on the sixth, seventh, or eighth day of the test. The maximum allowable test duration is 8 days. If properly stored and adequate in volume, the third sample may be used for renewal on the 8th day. Tests may not be continued beyond production of the third brood or past 7 days in order to get 15 neonates per surviving adult in the control.

Endpoints: number of survivors at seven days and number of neonates per female at three broods (# neonates per concentration divided by the # females at test initiation)

Control performance criteria: ≥ 80% survival in the control

an average of 15 neonates per surviving adult in the control

≥ 60 percent of the surviving control organisms producing three broods.

Other test acceptability criteria: ≤ 10% males in the surviving test organisms over all test concentrations

≤ 20% males in the surviving test organisms in the ACEC, CCEC, or LOEC

All surviving *Ceriodaphnia* producing no neonates in the test must be examined to determine gender, and the results of the determination reported. It is not necessary to identify gender when reproduction has been nearly eliminated in any test concentration when this fits an expected concentration-response relationship. It is understood that very young *Ceriodaphnia* can be difficult to sex and any *Ceriodaphnia* that dies in the first two days of the test may be excluded from calculations for reproduction if gender is difficult to determine and it is one of no more than two mortalities in a concentration. Otherwise, difficult to sex young *Ceriodaphnia* must be considered to be female and included in all calculations.

Fathead Minnow Survival and Growth

Test species: *Pimephales promelas*

Approved test method: EPA/600/4-91/002

Test type: 7-day static-renewal (80% renewal of test solution in each test chamber daily)

Temperature: 25° ± 1°C

Illumination: Illumination must be for 16 hours at 10 - 20 µE/m^2/s (50 - 100 ft-c) followed by 8 hours of darkness.

Test chamber size: 500 mL (minimum)

Test solution volume: 250 mL (minimum)

Age of test organisms: < 24 hours (< 48 hours if shipped)

Number of organisms/chamber: 10

Number of replicates/concentration: 4 (minimum)

Feeding: 0.1 g wet weight *Artemia* nauplii 3 times daily at 4 hour intervals (4 times/day at 2.5-3.0 hour intervals is acceptable) or 0.15 g wet weight *Artemia* nauplii twice daily at 6 hour intervals: no food in final twelve hours

Aeration: none unless DO < 4.0 mg/L; aerate all chambers and use < 100 bubbles/minute

Test duration: 7 days

Endpoints: the number of survivors and the total weight of survivors divided by the initial count (no zero weights except for reference toxicant testing)

Control performance criteria: ≥ 80% survival in the control

average dry weight ≥ 0.25 mg in the control

Selenastrum Growth

Test species: *Selenastrum capricornutum*

Approved test method: EPA/600/4-91/002

Test type: static (nonrenewal)

Temperature: 25° ± 1°C

Illumination: Illumination must be continuous at 86 ± 8.6 $\mu E/m^2/s$ (400 ± 40 ft-c or 4306 lux) and equally distributed over all test chambers.

Test chamber size: 125 mL or 250 mL

Test solution volume: for flasks shaken continuously - 50 mL test solution in 125 mL flasks or 100 mL test solution in 250 mL flasks

for flasks shaken twice daily by hand - 25 mL test solution in 125 mL flasks or 50 mL test solution in 250 mL flasks This option is not preferred and may be withdrawn.

Age of stocking solution: 4 to 7 days

Number of organisms/chamber: 10,000 cells/mL

Number of replicates/concentration: 4

Test duration: 96 hours

Endpoints: cell count only

Control performance criteria:

Controls must have at the end of the test 1,000,000 cells/mL with EDTA or 200,000 cells/mL without EDTA. The use of EDTA is not allowed unless special approval is granted because almost all effluents and receiving waters have the possibility of toxic concentrations of metals.

Variability of controls should not exceed 20% coefficient of variation.

Other test acceptability criteria:

A concurrent reference toxicant test must be conducted with each batch of tests.

3. Standard Saltwater Chronic Toxicity Tests

Permits for discharges to saltwater or brackish water will contain standard requirements for the use of a fish, topsmelt (*Atherinops affinis*) or silverside minnow (*Menidia beryllina*), and a mysid, *Holmesimysis costata* or *Mysidopsis bahia*, in chronic toxicity tests measuring survival and growth. New permits will instruct permittees to use the West Coast fish (topsmelt, *Atherinops affinis*) and mysid (*Holmesimysis costata*) for toxicity testing unless the lab cannot obtain a sufficient quantity of a West Coast species in good condition in which case the East Coast fish (silverside minnow,*Menidia beryllina*) or mysid (*Mysidopsis bahia*) may be substituted. Existing permits might contain a requirement for testing which only mentions the East Coast pair (*Menidia beryllina* and *Mysidopsis bahia*). However, we consider testing with the West Coast fish and mysid to be equivalent to the East Coast fish and mysid. If a lab wishes to minimize the transition period when testing will be done with organisms from both coasts, then the West Coast organisms can be tested in place of the East Coast organisms required in the permit. Labs should check with the client first because some permittees will want a letter from the Department of Ecology authorizing the switch. Tell cautious clients to write a letter to their Ecology facility manager requesting permission for the substitution.

The topsmelt and *Holmesimysis* tests are new to Washington state; labs needing assistance conducting the test or obtaining test organisms may call Brian Anderson or John Hunt of the University of California Marine Pollution Studies Lab at (408) 624-0947.

Labs do not need to attempt the fecundity endpoint with the mysid test. Success with the fecundity endpoint is too rare for it to have any use in the permitting program.

Labs can use brine in chronic toxicity testing with saltwater organisms, and the highest effluent concentration in the test will be around 70 percent.

All conditions in the following tables for the standard saltwater chronic toxicity tests must be met and reported for each test.

Holmesimysis Survival and Growth

Test species: *Holmesimysis costata*

Approved test method: EPA/600/R-95/136, August 1995

Test type: 7-day static-renewal (75% renewal of test solution in each chamber at 48 and 96 hours)

Temperature: 13° ± 1°C (No mysids allowed originating from south of Pt. Conception)

Illumination: Illumination must be for 16 hours at 10 - 20 µE/m^2/s (50 - 100 ft-c) followed by 8 hours of darkness.

Salinity: 30 or 34 ± 2‰

Test chamber size: 1000 mL (minimum)

Test solution volume: 200 mL (minimum)

Age of test organisms: 3 - 4 days post hatch

Number of organisms/chamber: 5

Number of replicates/concentration: 5 (minimum)

Feeding: twice daily (20 *Artemia* nauplii/mysid at each feeding); no food on day 7

Aeration: none unless DO < 4.0 mg/L; aerate all chambers and use < 100 bubbles/minute

Test duration: 7 days

Endpoints: the number of survivors and the total weight of survivors divided by the initial count (no zero weights except for reference toxicant testing)

Control performance criteria: ≥ 75% survival in the control

average dry weight ≥ 0.40 mg in the control

Reference toxicant acceptability criteria: MSD < 40% (survival) and 50 µg (growth)

survival and growth NOECs < 100 µg/L in a zinc sulfate reference toxicant test.

Mysidopsis Survival and Growth

Test species: *Mysidopsis bahia*

Approved test method: EPA/600/4-91/003

Test type: 7-day static-renewal (90% renewal of test solution in each test chamber daily)

Temperature: 26° ± 1°C

Illumination: Illumination must be for 16 hours at 10 - 20 µE/m^2/s (50 - 100 ft-c) followed by 8 hours of darkness.

Salinity: 30 ± 2‰

Test chamber size: 8 oz plastic disposable cups or 400 mL glass beakers (minimum)

Test solution volume: 150 mL (minimum)

Age of test organisms: 7 days

Number of organisms/chamber: 5

Number of replicates/concentration: 8 (minimum)

Feeding: twice daily (75 *Artemia* nauplii/mysid at each feeding) with 8 - 12 hours between feedings

Aeration: none unless DO < 4.0 mg/L; aerate all chambers and use < 100 bubbles/minute

Test duration: 7 days

Endpoints: the number of survivors and the total weight of survivors divided by the initial count (no zero weights except for reference toxicant testing)

Control performance criteria: ≥ 80% survival in the control

average dry weight ≥ 0.20 mg in the control

Topsmelt Survival and Growth

Test species: *Atherinops affinis*

Approved test method: EPA/600/R-95/136

Test type: 7-day static-renewal (75% renewal of test solution in each test chamber daily)

Temperature: 20° ± 1°C

Illumination: Illumination must be for 16 hours at 10 - 20 µE/m^2/s (50 - 100 ft-c) followed by 8 hours of darkness.

Salinity: 30 or 34 ± 2‰

Test chamber size: 600 mL (minimum)

Test solution volume: 200 mL (minimum)

Age of test organisms: 9 - 15 days post-hatch

Number of organisms/chamber: 5

Number of replicates/concentration: 5 (minimum)

Feeding: twice daily (40 *Artemia* nauplii/fish at each feeding) morning and afternoon; no food on day 7.

Aeration: none unless DO < 4.0 mg/L; aerate all chambers and use < 100 bubbles/minute

Test duration: 7 days

Endpoints: the number of survivors and the total weight of survivors divided by the initial count (no zero weights except for reference toxicant testing)

Control performance criteria: ≥ 80% survival in the control; average dry weight ≥ 0.85 mg in the control

Reference toxicant acceptability criteria: MSD < 25% (survival) and 50% (growth)

LC_{50} < 205 µg/L in a copper chloride reference toxicant test.

Inland Silverside Survival and Growth

Test species: *Menidia beryllina*

Approved test method: EPA/600/4-91/003

Test type: 7-day static-renewal (80% renewal of test solution in each test chamber daily)

Temperature: 25° ± 1°C

Illumination: Illumination must be for 16 hours at 10 - 20 µE/m^2/s (50 - 100 ft-c) followed by 8 hours of darkness.

Salinity: 30 ± 2‰

Test chamber size: 600 - 1000 mL

Test solution volume: 500 - 750 mL

Age of test organisms: 7 - 11 days

Number of organisms/chamber: 10 - 15 as long as each test chamber contains the same number and test chamber sizes and test solution volumes toward the larger end of the acceptable range are used for larger numbers of fish

Number of replicates/concentration: 4

Feeding: 0.10 g wet weight *Artemia* nauplii once per day per replicate through day 2; 0.15 g wet weight per replicate on days 3 - 6; no food on day 7

Aeration: none unless DO < 4.0 mg/L; aerate all chambers and use < 100 bubbles/minute

Test duration: 7 days

Endpoints: the number of survivors and the total weight of survivors divided by the initial count (no zero weights except for reference toxicant testing)

Control performance criteria: ≥ 80% survival in the control

average dry weight ≥ 0.50 mg in the control

4. Supplemental Saltwater Chronic Toxicity Tests

Permits for discharges to saltwater might include one of the following supplemental saltwater chronic toxicity tests.

The bivalve embryo-larval development test will be placed into a permit along with the standard fish and invertebrate test when there is a risk of toxicity to sensitive larval life-stages of marine organisms. This test is especially appropriate for discharges to ecosystems of special importance or fragility which are breeding grounds for marine organisms. The bivalve test is also appropriate for discharges to inlets or bays with poor circulation or for larger discharges with a tendency to stratify. The echinoderm development test is a potential alternative to the bivalve development test.

The combination of sensitivity with very short duration is unique to the echinoderm fertilization test. Very small volumes of effluent can be tested successfully and one spawning yields enough material for many tests. The echinoderm fertilization test will be included in a permit when a balance between high sensitivity and convenience are important.

If the receiving water contains or should contain kelp beds (shallow and rocky), then the *Macrocystis* germination and growth test might be required. If an effluent is suspected to be phytotoxic, then the *Macrocystis* test might also be required. The *Macrocystis* test is new to Washington state; labs needing assistance conducting the test or obtaining test organisms may call Brian Anderson or John Hunt of the University of California Marine Pollution Studies Lab at (408) 624-0947.

All conditions in the following tables for the supplemental saltwater chronic toxicity tests must be met and reported for each test.

Bivalve Development

Test species: *Crassostrea gigas* or *Mytilus sp.* (*M. trossulus, M. galloprovincialis, M. californianus*)

Approved test method: EPA/600/R-95/136

Test type: static (nonrenewal)

Temperature: 20° ± 1°C for oysters, 15° or 18° ± 1°C (16° ± 1° if already the lab's standard temperature) for mussels

Illumination: Illumination must be for 16 hours at 10 - 20 µE/m^2/s (50 - 100 ft-c) followed by 8 hours of darkness.

Salinity: 30 ± 2‰

Test chamber size: 30 mL

Test solution volume: 10 mL

Age of test organisms: < 4 hours after fertilization

Number of organisms/chamber: 150 - 300

Number of replicates/concentration: 4

Aeration: none in test chambers; the sample may be aerated if the DO < 4.0 mg/L

Test duration: 48 hours (up to 54 hours in order to achieve complete development)

Endpoints:

1. Calculate the EC_{25} (or EC_{50} if Probit cannot be used) for proportion normal and for proportion alive.

2. If the EC_{25} or EC_{50} for proportion alive is less than the same point estimate calculated for proportion normal or if the 95% confidence limits overlap, then calculate a combined proportion normal/alive and use it as the test endpoint. Otherwise, use the proportion normal as the test endpoint.

3. If a combined proportion normal/alive is used and proportions greater than 1.0 occur, then the number normal must be used for any hypothesis testing performed on the test data.

For more discussion of the calculation of the bivalve development endpoint, see Appendix B.

Test acceptability criteria:

Bivalve development tests will be evaluated for compliance with the following test acceptability criteria rather than the list in item 16 in Table 4 of the EPA manual. The test will be reviewed for compliance with all other conditions and procedures specified in the EPA manual and in section 13 of ASTM E 724.

A test is acceptable if ≥ 70% of oyster or mussel embryos introduced into the dilution water control grew into live larvae with completely developed shells at the end of the test.

A test is acceptable if the minimum significant difference is < 25%.

Unless all embryos are counted in each test chamber at the beginning of the test to get a true start count, the estimated initial count is derived from the mean of the counts of at least 6 extra test chambers prepared exactly as the control test chambers using a procedure that randomly distributes their preparation throughout the setting up of all the test chambers.

The coefficient of variation should be ≤ 15% for the embryo counts on the minimum of 6 subsamples taken from the stocking solution at the beginning of the test in order to estimate an initial count. If the 15% coefficient of variation is exceeded, the test report must note this fact and warn to use the test result with caution. Tests will not be rejected solely for exceeding the 15% coefficient of variation.

A concurrent reference toxicant test must be conducted with each batch of tests.

Echinoderm Fertilization

Test species: *Strongylocentrotus purpuratus* or *Dendraster excentricus*

Approved test method: EPA/600/R-95/136

Test type: static (nonrenewal)

Temperature: 12° ± 1°C

Salinity: 30 ± 2‰

Test chamber size: 16 × 100 mm or 16 × 125 mm disposable culture tubes

Test solution volume: 5 mL

Age of test organisms: < 4 hours after collection of gametes

Number of spawners: Gametes are pooled from ≤ 4 males and ≤ 4 females (≤ 6 female sand dollars)

Number of organisms/chamber: Approximately 1,120 eggs and ≤ 3,360,000 sperm

Number of replicates/concentration: 4

Aeration: none in test chambers; the sample may be aerated if the DO < 4.0 mg/L

Test duration: 40 minutes (20 minutes exposure of sperm; 20 minutes with eggs)

Endpoints: fertilization of eggs (elevation of the fertilization membrane)

Test acceptability criteria:

A test is acceptable if ≥ 70% of eggs in the control are fertilized.

A test is acceptable if the minimum significant difference is < 25%.

Fertilization at the NOEC must be within 80% of control fertilization.

A concurrent reference toxicant test must be conducted with each batch of tests.

Dilution water egg blanks and effluent egg blanks should contain essentially no eggs with fertilization membranes.

The sperm count for the final sperm stock must be≤ 33,600,000/mL and one of the following options met:

Option 1, trial fertilization used - The sperm count for the final sperm stock must not exceed double the target density determined from the fertilization trial test used to determine the sperm density that will provide about 80% to 100% fertilization without oversperming.

Option 2, sperm/egg ratio kept ≤ 500:1 - confirmation of a sperm stock density of ≤ 5,600,000/mL

Option 3, use any reasonable sperm stock density and run two extra sets of controls (a high and a low density control) - the high density control (0.2 mL sperm stock) must have at least 5% higher fertilization than the low density control (0.05 mL sperm stock).

Echinoderm Development

Test species: *Strongylocentrotus purpuratus* or *Dendraster excentricus*

Approved test method: EPA/600/R-95/136

Test type: static (nonrenewal)

Temperature: 15° ± 1°C

Illumination: Illumination must be for 16 hours at 10 - 20 µE/m^2/s (50 - 100 ft-c) followed by 8 hours of darkness.

Salinity: 30 ± 2‰

Test chamber size: 30 mL

Test solution volume: 10 mL

Age of test organisms: ≤ 1 hour after fertilization

Number of organisms/chamber: Approximately 250 fertilized eggs in 0.25 mL of egg solution

Number of replicates/concentration: 4

Aeration: none in test chambers; the sample may be aerated if the DO < 4.0 mg/L

Test duration: 72 hours

Endpoints:

1. Calculate the EC_{25} (or EC_{50} if Probit cannot be used) for proportion normal and for proportion alive.

2. If the EC_{25} or EC_{50} for proportion alive is less than the same point estimate calculated for proportion normal or if the 95% confidence limits overlap, then calculate a combined proportion normal/alive and use it as the test endpoint. Otherwise, use the proportion normal as the test endpoint.

3. If a combined proportion normal/alive is used and proportions greater than 1.0 occur, then the number normal must be used for any hypothesis testing performed on the test data.

The endpoint of the echinoderm development test should be the same as the endpoint for the bivalve development test. For a discussion of the calculation of the bivalve development endpoint, see Appendix B.

Test acceptability criteria:

A test is acceptable if ≥ 80% of larvae in the control have developed normally.

A test is acceptable if the minimum significant difference is < 25%.

Unless all embryos are counted in each test chamber at the beginning of the test to get a true start count, the estimated initial count is derived from the mean of the counts of at least 6 extra test chambers prepared exactly as the control test chambers using a procedure that randomly distributes their preparation throughout the setting up of all the test chambers.

The coefficient of variation should be ≤ 15% for the embryo counts on the minimum of 6 subsamples taken from the stocking solution at the beginning of the test in order to estimate an initial count. If the 15% coefficient of variation is exceeded, the test report must note this fact and warn to use the test result with caution. Tests will not be rejected solely for exceeding the 15% coefficient of variation.

A concurrent reference toxicant test must be conducted with each batch of tests.

Macrocystis Germination and Growth

Test species: *Macrocystis pyrifera*

Approved test method: EPA/600/R-95/136

Test type: static (nonrenewal)

Temperature: 15° ± 1°C

Illumination: Illumination must be for 16 hours at 50± 10 µE/m^2/s equally distributed over all test chambers followed by 8 hours of darkness.

Salinity: 34 ± 2‰

Test chamber size: 600 mL

Test solution volume: 200 mL

Age of test organisms: < 2.5 hours after sporophylls begin releasing zoospores

Number of organisms/chamber: 7,500 zoospores/mL

Number of replicates/concentration: 5

Aeration: none unless DO < 4.0 mg/L; aerate all chambers and use < 100 bubbles/minute.

Test duration: 48 hours

Endpoints: Percent of zoospores with germination tubes at least one spore diameter in length

Average length of 10 germination tubes randomly selected from each test chamber

Test acceptability criteria: ≥ 70% germination of zoospores in the control

≥ 10 µm average germ tube length in the control

Reference toxicant acceptability criteria: NOEC < 35 µg/L in a concurrent copper chloride reference toxicant test.

The MSD is < 20% relative to the control for both germination and germ tube length in the copper chloride reference toxicant test.

Appendix A
Rainbow Trout Age Discussion

The Department of Ecology's intent is to evaluate WET tests consistently in accordance with EPA protocols. The purpose of fish age criteria is to standardize testing to a sensitive stage of the fish's life cycle. We are concerned that the age of rainbow trout is being determined differently from lab to lab because the point of the fish's life cycle representing day 1 is not always the same.

The EPA protocol for the acute rainbow trout test sets an age requirement for the fish of 15 to 30 days old. There has been some uncertainty, however, at what point in the life cycle is day 1. This issue was researched through consultations with fish biologists, labs, and EPA. Little agreement exists about the upper end of the sensitive age range for rainbow trout testing, and many believe that EPA might be too restrictive on the upper age. There is general agreement, however, that testing should not begin until after the yolk sac is completely absorbed and the fish are actively feeding. Swim-up is believed to be the least ambiguous event to use in timing the readiness of trout for testing.

In accordance with the findings of these consultations, Ecology intends to evaluate rainbow trout acute test fish age criteria as follows:

- Ecology will enforce the EPA age range of 15 to 30 days old. Fish age will be determined using swim-up as day 1. Labs must express rainbow trout age as days after swim-up.

- The fish should be held at 12±1°C after reaching the swim-up life stage. This ensures that fish age and condition are consistent.

The test fish should be the same age and from the same source. Because of individual development rate variation, test fish will be considered to be at a stage in their life cycle when 80% of the batch have achieved that stage. Rainbow trout development is temperature dependent. 12°C is the preferred rearing temperature, but trout may be held at a lower temperature prior to swim-up.

The life cycle stage definitions are:

Hatch: When the fish (alevins) have broken out of the egg casing, but are inactive, remain mostly on the bottom, do not feed, and live off the attached yolk sac.

Swim-up: Around 3 weeks from hatch, the fish emerge from the relatively inactive bottom dwelling stage and actively move up and remain in the upper water column. The fish have begun feeding but still have some yolk sac.

NOTE: Because of occasional shortages of rainbow trout of the correct age for testing, we will begin accepting brook trout (*Salvelinus fontinalis*) as a substitute using the same test conditions for rainbow trout except that the age range is 30 to 60 days post-hatch and at least 2 days past swim-up. Be sure to check to get your client's OK before making the substitution.

Appendix B
Bivalve Development Test Endpoint Discussion

A. INTRODUCTION

On March 4, 1996, a meeting of scientists familiar with the bivalve embryo-larval development test was held in Portland, Oregon to discuss issues involving the test endpoints. The meeting discussions focused on two main questions involving the choice of endpoint calculation. Which endpoints are preferred based on variability and which endpoints are preferred based on scientific considerations? The meeting attendees decided, based on data from the State of Washington variability study, that the recommendation of the Biomonitoring Science Advisory Board (BSAB) in favor of the bivalve development test based on the variability of the proportion normal endpoint would not be changed for proportion normal/alive (combined endpoint).

The EPA 1995 bivalve test contains an adjusted combined normal/alive proportion calculation where the # normal for each replicate is divided by the larger of the initial or final count. Because the initial count is based on a mean of the counts on subsamples, the final count or # normal for some replicates will sometimes exceed the initial count. The EPA adjustment avoids the generation of proportions greater than 1 and is also an attempt to increase test sensitivity. The adjustment was determined by the group to be unnecessary to increase test sensitivity. The bivalve development test is already very sensitive and data indicates that the adjusted combined endpoint does little to increase sensitivity anyway.

The adjusted combined endpoint calculation introduces bias and complicates hypothesis testing. If the final count is greater than the initial count, it is assumed to be due to subsampling differences and the final count is used in the denominator. However, the calculation implies that toxicity is always the cause for initial counts being greater than final counts even though final counts will sometimes be greater than initial counts due to variability alone when the initial count is based on the mean of the counts on several subsamples. This situation may also violate the independence of observation assumption required for valid parametric hypothesis testing procedures. After consideration of these circumstances, the group decided to recommended against the use of the adjusted combined endpoint in the EPA manual.

In addition, the attendees developed a process for determining which endpoint, proportion normal or proportion normal/alive, to use for the results of any bivalve development test. This process is described in detail below. The only change from the process recommended at the meeting is the use of the EC_{25} or EC_{50} instead of the NOEC for comparing the sensitivity of the endpoints. Point estimates such as the EC_{25} or EC_{50} are better than the NOEC for comparisons between tests, and because of the possibility of proportions greater than 1, valid NOECs will not always be available for use in the process. The 95% confidence limits for the point estimates are useful in comparisons because data have shown that mortalities can have a significant effect on the proportion normal/alive even when proportion alive is not the most sensitive endpoint.

The attendees also recommended combining the separate control performance criteria for survival and for development in the EPA West Coast manual into a normal/alive control performance criterion that is similar to that in ASTM and PTI '94. The control performance criterion for mussels was to be raised to

equal that for oysters if Washington Department of Ecology data indicated that the higher performance was a reasonable expectation. Data indicate that mussel controls perform as well as oyster controls.

The attendees recommended that the initial count be determined from the mean of the counts from at least 6 extra test chambers prepared exactly as the control test chambers using a procedure that randomly distributes their preparation throughout the setting up of all the test chambers, and that a warning level of 15% coefficient of variation be applied to the counts on these test chambers. A coefficient of variation ≤ 15% will mean that not only is the initial count reasonably accurate, but that lab pipetting and counting technique are generally good.

B. ENDPOINT CALCULATION PROCESS

The proportion normal is the preferred endpoint unless the test has significant mortality in which case the combined proportion normal/alive is the preferred endpoint. To determine the preferred endpoint for a test conduct the following:

1. Calculate the EC_{25} (or EC_{50} if Probit cannot be used) for proportion normal and proportion alive.

2. If the EC_{25} or EC_{50} for proportion alive is less than the same point estimate calculated for proportion normal or if the 95% confidence limits overlap, then calculate a combined proportion normal/alive to use as the test endpoint. Otherwise, use proportion normal as the test endpoint.

3. If a combined proportion normal/alive is used and proportions greater than 1.0 occur, then the number normal must be used for any hypothesis testing performed on the test data.

C. TERMINOLOGY AND EQUATIONS

initial count = the mean of a minimum of 6 subsamples taken from the stocking solution

normal = number of larvae at the end of the test with completely developed shells*

abnormal = number of larvae at the end of the test with incompletely developed shells*

final count = # normal + # abnormal

proportion alive = final count ÷ initial count

proportion normal = # completely developed ÷ final count

combined proportion normal/alive = # completely developed ÷ initial count

* See the test method for a more complete description.

D. TEST ACCEPTABILITY CRITERIA DECISIONS

A test is acceptable if≥ 70% of oyster or mussel embryos introduced into the dilution water control grew into live larvae with completely developed shells at the end of the test.

Unless all embryos are counted in each test chamber at the beginning of the test to get a true start count, the estimated initial count is derived from the mean of the counts of at least 6 extra test chambers prepared exactly as the control test chambers using a procedure that randomly distributes their preparation throughout the setting up of all the test chambers. These extra chambers will be used at the beginning of the test in order to estimate an initial count and assess pipetting and counting technique. The coefficient of variation must be≤ 15% for the embryo counts on these subsamples. If the 15% coefficient of variation is exceeded, the test report must warn to use the test result with caution. Tests will not be rejected solely for exceeding the 15% coefficient of variation.

Appendix C
Growth or Combined Survival and Growth Endpoint Discussion

EPA changed the growth calculation for the 7-day survival and growth tests in the new chronic toxicity testing manuals referenced in this document. Instead of dividing the final weight by the number of surviving organisms at the end of the test, the new chronic manuals instruct the lab to divide by the number of organisms at test initiation. The new endpoint calculation results in a combined survival and growth number.

If all of the test organisms survive, then the original growth calculation and the combined survival and growth calculation result in the same numbers. If an effluent produces significant mortality with a steep concentration-response, then the NOEC for the test tends to be the same for the original proportion alive and the combined survival and growth endpoint. If there are partial mortalities at effluent concentrations below the LOEC for proportion alive, the combined survival and growth calculation will increase test organism response relative to the original growth calculation, but it will also increase variability across the replicates as well. The increased variability decreases statistical sensitivity resulting in about equal sensitivity for the original growth and the combined survival and growth endpoints. Published EPA data show no increased test sensitivity from the combined survival and growth endpoint using fathead minnow (See **Pickering, Q., J. Lazorchak** and **K. Winks**. 1996. Subchronic sensitivity of one-, four-, and seven-day old fathead minnow (*Pimephales promelas*) larvae to five toxicants. *Environ. Toxicol. Chem.* 15:353-359.) Department of Ecology data on the 7-day survival and growth tests using three different species of test organisms also show no increased sensitivity from changing the endpoint calculation and an increased tendency toward anomalous tests as described in Appendix D.

The Department of Ecology WET database has shown that the combined endpoint for mortality/weight has greater variability than the original growth endpoint and often shows both an increased apparent effect and reduced statistical sensitivity. If there are control mortalities (the EPA manuals allow tests that have as low as 80% survival in the control), then the apparent toxic effect can be smaller than with the original growth calculation. These consequences tend to cancel one another resulting in little difference in test outcome overall from the original endpoint.

In order to not be too far out of line with other states and because EPA argues in favor of the combined endpoint, we will make the change and accept the increased test variability with the combined endpoint. However, when sporadic mortalities occur, the variability becomes unacceptable. Therefore, tests that have a standard deviation for proportion alive above 0.25 in any effluent concentration (unless the partial mortality occurs at the threshold of toxicity in a good concentration-response relationship) will be analyzed for the original growth endpoint.

The need for switching back to the original growth calculation when survival is highly variable can sometimes be avoided by not using zero weights. Zero weights make no sense for the original growth calculation (weight/final count) since zero weights can only happen if everything died and 0/0 is undefined. Zero weights are also not practical for the combined survival and growth calculation (weight/initial count). It is true that there is zero biomass when everything dies, but if this occurs to nearly the same degree in every test chamber at that concentration, then that concentration will certainly have a statistically significant reduction in survival and the result for combined survival and growth will

be superfluous. However, if everything dies in a test chamber and survival is fairly good in other test chambers at the same concentration, then the zero weight for the one test chamber can cause a high standard deviation and little statistical sensitivity making the low mean weight for that concentration irrelevant to the test results. Anomalous concentration-response relationships will occur when a zero weight in one replicate reduces the mean weight for that concentration enough to overcome low statistical sensitivity while survival and growth are generally good in other test chambers and concentrations. When zero weights work well with the combined survival and growth calculation, the results are superfluous because survival by itself is enough. When zero weights don't work well with the combined survival and growth calculation, statistical sensitivity or concentration-response suffers. For these reasons, zero weights are not used (space is left blank) with either weight calculation here in Washington State. The only exception would be reference toxicant testing where an IC_{25} is needed for control charting.

Only enter weights when something is weighed. Weight is a property of mass. If there are no test organisms left, the weight is not zero but meaningless.

Appendix D
Identifying Anomalous WET Tests

Introduction

These guidelines are intended to supplement Chapter 173-205 WAC (the WET rule) in defining anomalous WET test results. WAC 173-205-070(5)(c) states that anomalous WET test results will be identified and not used for compliance determinations. WAC 173-205-090(1)(d) describes the process for a permittee to notify Ecology that noncompliance with a WET limit may have been caused by an anomalous WET test result. If a WET test result indicates noncompliance with a WET limit but will be identified later by Ecology as anomalous, a permittee can avoid the expense of unnecessary extra WET testing by submitting notification of an anomalous WET test result to Ecology. The notification must include the reason for considering the test result to be anomalous. If Ecology agrees with the permittee's reason for considering the test result to be anomalous, the additional monitoring required by WAC 173-205-090(1) will be avoided. A list of criteria at the end of these guidelines contains some of the considerations that Ecology will use in deciding if WET test results are anomalous.

Text of WAC 173-205-090(1)(D)

WAC 173-205-090(1)(d) If the permittee believes that the compliance test failure will be identified by the Department (Ecology) as an anomalous test result in accordance with WAC 173-205-070(5)(c), the permittee may send the Department notification with the compliance test result that the compliance test result might be anomalous and that the permittee intends to take only one additional sample for toxicity testing and wait for notification from the Department before completing the additional monitoring required in this subsection.

(i) The notification must identify the reason for considering the compliance test result to be anomalous.
(ii) The permittee shall take the additional sample and retest as soon as possible after receiving the compliance test result.
(iii) The additional test result shall replace the compliance test result upon determination by the Department that the compliance test result was anomalous.
(iv) The permittee shall complete all of the additional monitoring required by this subsection as soon as possible after notification by the Department that the compliance test result was not anomalous.
(v) If the additional sample fails the compliance test, then the permittee shall proceed without delay to complete all of the additional monitoring required by this subsection.

The Difference Between Invalid Tests and Anomalous Test Results

Invalid WET tests occur when the lab does not follow the test protocol or when the results do not meet the test acceptability criteria in the test protocol. Permittees and labs are obligated to look for invalid tests because the permit requires that the test protocol be followed. Ecology will also be reviewing WET test results to see that they are based on valid tests.

Anomalous test results happen when the lab appears to have conducted the WET test in accordance with the test protocol, but the results are considered unreliable according to the following anomalous test identification criteria. There is no requirement for permittees to attempt to identify anomalous WET test results, and all valid WET test results must be submitted whether the test is regarded as anomalous or not. Ecology will be reviewing all WET test results to identify invalid tests and anomalous test results. The anomalous test identification criteria, listed below for the use of permittees and labs, will also guide Ecology in identifying anomalous WET test results. The identification of an anomalous test result does not by itself imply any fault on the part of the permittee or lab, but frequent anomalous tests can be an indication of poor lab technique or poor condition of test organisms.

The main purpose for conducting effluent toxicity tests with at least five effluent concentrations in a series is to allow concentration-response to be evaluated and anomalous tests discarded. The identification of anomalous tests is a valuable tool for reducing false positives. A concentration-response relationship where response increases with concentration is a good identifier of toxicity as opposed to other sources of organism stress such as disease. Test method variability or lab error will also very rarely produce a good concentration-response relationship. Identifying a test as anomalous does not necessarily mean rejection of the test and a requirement to repeat. If a test result meets one of the criteria for anomalous test identification but has no statistically significant toxicity at concentrations of regulatory concern (ACEC or CCEC), then the test need not be repeated unless other factors contribute to a decision to reject the test.

The anomalous test identification criteria are a common sense approach to making WET test results fair and enforceable. They should be taken at face value and are not intended to have defined statistical confidence levels or rely on sophisticated curve-fitting models. The anomalous test criteria will be used during test review to intervene with human judgment when statistics seem to be reaching the wrong conclusion about effluent toxicity. Their underlying principle is the definition of the NOEC as the highest effluent concentration showing no statistically significant difference from the control along with an expectation for a concentration-response relationship typical for toxicity under the conditions of the test.

Different toxicity tests have different expectations for a good concentration-response relationship. The proportional endpoints (survival, echinoderm fertilization, bivalve development) have steeper concentration-response relationships than do the nonproportional endpoints such as growth or neonate production. Some bivalve development tests have two distinct stepwise effect thresholds, a development effect threshold followed by a survival effect threshold at a higher concentration. Water chemistry gradients will sometimes modify the expected concentration-response relationship. The anomalous test definitions must be considered in light of the expectations for the different toxicity tests and endpoints.

Notification of an Anomalous Test Result

When a WET test result does not comply with a WET limit, the permittee is required to begin additional monitoring as soon as possible. If the noncompliance was with an acute WET limit, additional monitoring is conducted weekly for four weeks. If the noncompliance was with a chronic WET limit, additional monitoring is conducted monthly for three months.

The WET rule allows a permittee to avoid the cost of the additional monitoring when noncompliance with a WET limit is believed to be due to an anomalous WET test result. A good laboratory will be able to inform a permittee of a likely anomalous WET test result that resulted in noncompliance with a WET limit. A permittee can then send Ecology notification with the compliance test result that the test might be anomalous and that the permittee intends to take only one additional sample for toxicity testing. If the additional sample fails to comply with the WET limit, then the permittee must proceed without delay to complete all of the additional monitoring. Otherwise, the permittee is not required to conduct the rest of the additional monitoring unless Ecology determines that the WET test result was not anomalous. The additional test result replaces the compliance test result upon determination by Ecology that the compliance test result was anomalous.

A permittee benefits from notifying Ecology of an anomalous test result only when there is noncompliance with a WET limit. The notification allows the permittee to delay the additional monitoring required after a WET limit violation while Ecology evaluates the notification and test result. The notification will also help Ecology determine sooner that the test result is anomalous and does not represent a WET limit violation. However, permittees that notify Ecology of anomalous test results that comply with WET limits would be duplicating Ecology's efforts with no benefit to themselves.

Permittees should exercise judgment about notification of anomalous WET test results. The WET rule gives Ecology the authority to determine which test results are anomalous, and Ecology may reject any permittee notification that does not meet the anomalous test identification criteria. Frequent anomalous test results will not be an effective shield against WET limit violations because they are likely to cause increased scrutiny of the permittee and the lab.

Resampling After Anomalous Test Result Identification

In order to satisfy a permit requirement for compliance monitoring, an anomalous test result must be replaced by a WET test result that can be used for compliance determinations. WAC 173-205-090(1)(d)(ii) requires a permittee to resample as soon as possible and conduct another WET test as part of the process of notifying Ecology of an anomalous WET test result. The permittee must also resample and conduct another WET test after being notified by Ecology of an anomalous test result. The cost of the repeated sampling and testing will be another disincentive to frequent anomalous test results.

Criteria for Identifying Anomalous Test Results

1. A WET test result is anomalous if it shows a statistically significant difference in response between the control and the ACEC or CCEC, but no statistically significant difference in response at one or more higher effluent concentrations. The lack of statistical significance must be associated with a lower toxic effect at the higher effluent concentration. Any higher effluent concentration used in this determination must be a part of a dilution series. Labs should not cluster test concentrations just above the ACEC or CCEC in order to increase the opportunity for an anomalous test result.

2. A WET test is anomalous if there is a statistically significant difference in response between the control and the ACEC or CCEC and the slope of the line fitted to the concentration-response plot of all test concentrations is zero, unless the zero slope is due to a complete effect (no survival, no fertilization, no normal development, etc.) at every effluent concentration.

3. A WET test is anomalous if there is a statistically significant difference in response between the control and the ACEC or CCEC which together with other nearby concentrations of effluent have a zero slope and appear to be nontoxic (performance is typical of healthy test organisms). Another description of this criterion is a test with a control that seems to not belong to the concentration-response relationship because of exceptionally good performance.

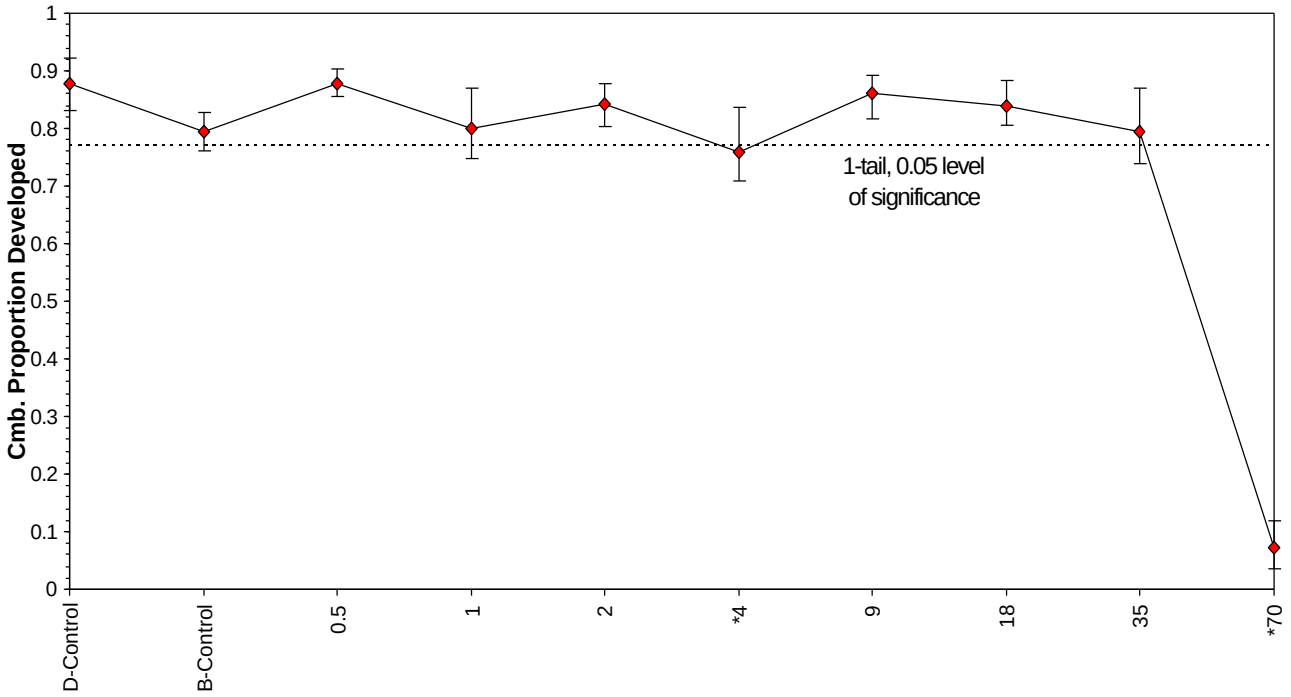
4. A WET test is anomalous if the overall slope of the line fitted to the concentration-response plot is opposite of normal expectations and there is a statistically significant difference in response at the ACEC or CCEC. A test might be considered acceptable if the slope is opposite over only part of the concentration series.

5. A WET test is anomalous if the standard deviation for proportion alive equals or exceeds 0.3 in any test concentration unless the partial mortality fits a good concentration-response relationship. A WET test is anomalous if mortalities occur in any test concentration in excess of the control performance criterion for survival when the concentration-response relationship indicates that the effluent concentration is nontoxic (sporadic mortalities).

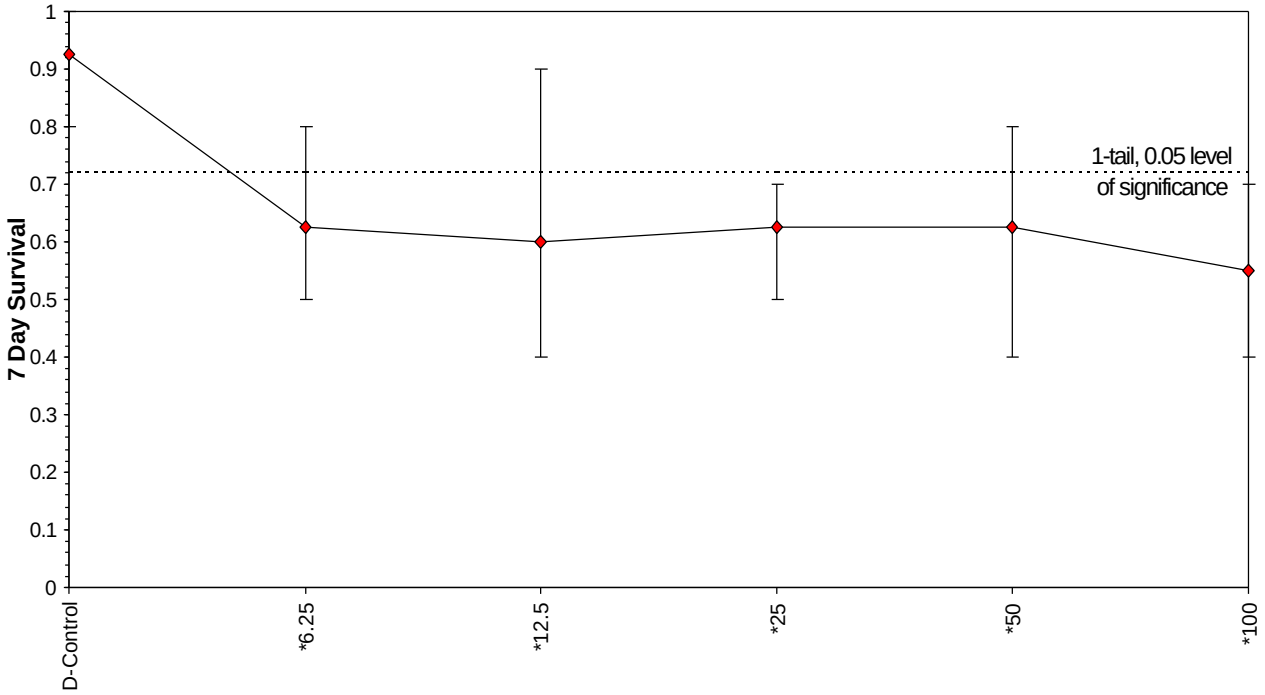
6. To reduce the opportunity for WET limit violations due to statistically significant differences in response that are type I errors, permit requirements will lower the alpha level for hypothesis testing when differences in test organism response are small. To prevent excessive type I errors, eliminate some interrupted concentration-response relationships, and have more fair and enforceable test results, we will setalpha = 0.01 for small differences in response. If the difference in survival between the control and the ACEC in an acute test is less than 10 percent, the level of significance will be lowered from 0.05 to 0.01. If the difference in test organism response between the control and the CCEC in a chronic test is less than 20 percent, the level of significance will be lowered from 0.05 to 0.01.

   If a permit with a WET limit does not specify this change in level of significance and differences in response are less than 10 percent (acute) or 20 percent (chronic), the lab should conduct the hypothesis test at both levels of significance. The permittee should report any discrepancy between the results at the two levels of significance as an anomalous test result.

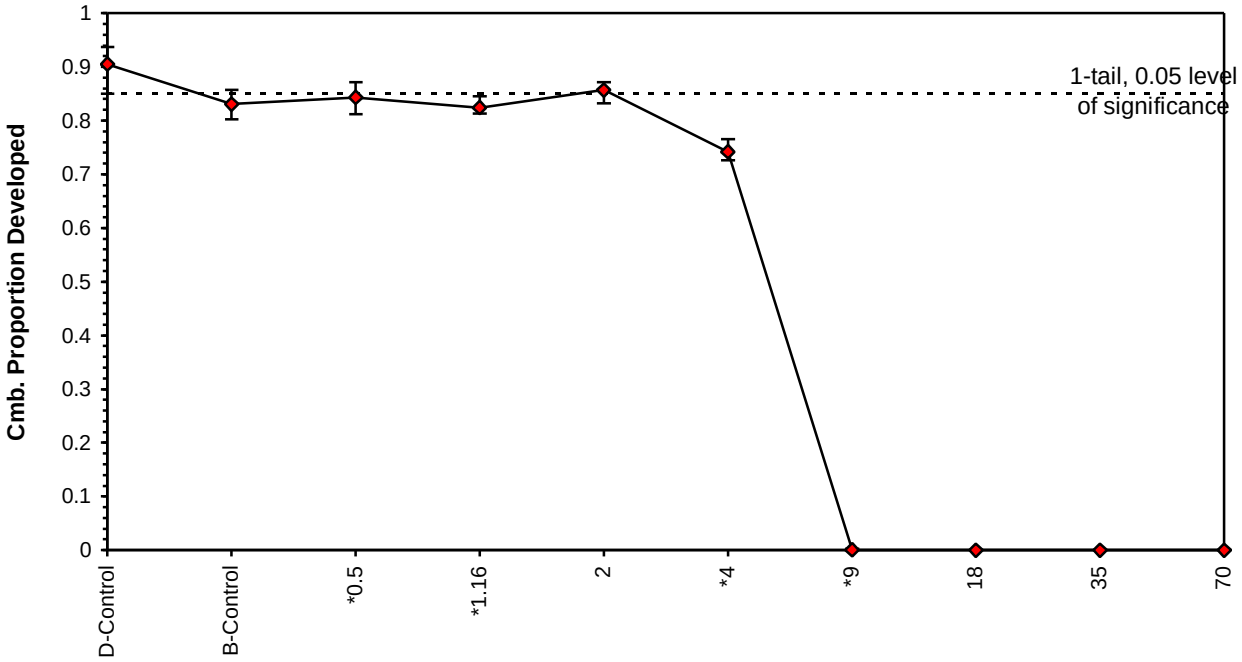
Example Test for Anomalous Test Criterion 1
bivalve development test, AQTX0993, on an industrial effluent



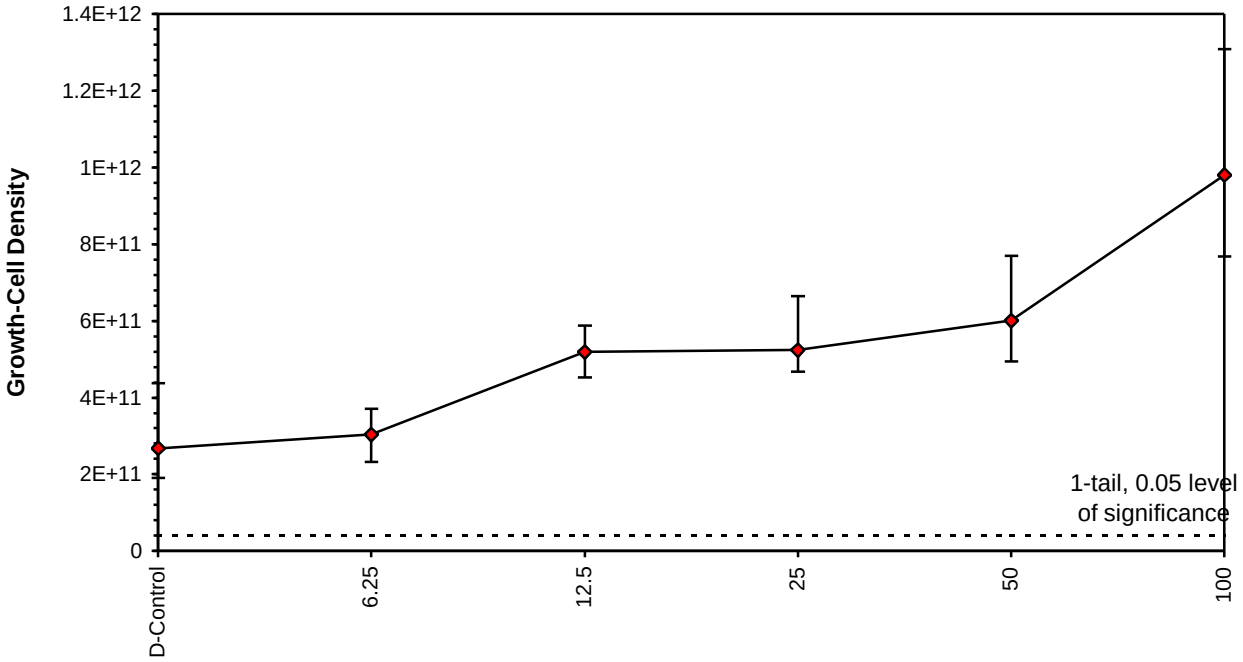
Example Test for Anomalous Test Criterion 2 (also matches criterion 3)
fathead minnow chronic test, KJOI356, on an industrial effluent



Example Test for Anomalous Test Criterion 3 (also matches criterion 1)
Selenastrum test, KJOI201, on an industrial effluent



Example Test for Anomalous Test Criterion 4
bivalve development test, AQTX0996, on an industrial effluent



Appendix E
Example Calculations for the Power Standards

ACEC	Fathead minnow- number surviving				
	replicate 1	replicate 2	Replicate 3	replicate 4	mean of replicates
25% effluent	6	4	8	7	6.25
Control	Fathead minnow- number surviving				
	replicate 1	replicate 2	Replicate 3	replicate 4	mean of replicates
lab water	9	10	9	9	9.25

1. Subtract the mean survival across the replicates in the ACEC from the mean survival across the replicates in the control.

$$9.25 - 6.25 = 3.00$$

2. Divide this difference between the mean survivals by the mean survival across the control replicates.

$$3.00 \div 9.25 = 0.32$$

3. Multiply the result by 100 and express as a percent difference in survival.

$$0.32 \times 100 = 32\% \text{ difference in response}$$

4. If the percent difference in survival is≤ 29%, then the WET test has met the power standard.

The 32% difference in response is > 29%

The WET test has not met the power standard and must be repeated. (Assuming that the WET test did not violate the WET limit; the power standards are not an issue for WET tests that violate WET limits.)

CCEC	Fathead minnow- average weight/larva (mg)				
	replicate 1	replicate 2	Replicate 3	replicate 4	mean of replicates
5% effluent	0.529	0.554	0.425	0.373	0.470
Control	Fathead minnow- average weight/larva (mg)				
	replicate 1	replicate 2	Replicate 3	replicate 4	mean of replicates
lab water	0.560	0.636	0.613	0.452	0.565

1. Subtract the mean of the responses across the replicates in the CCEC from the mean of the responses across the replicates in the control.

$$0.565 - 0.470 = 0.095$$

2. Divide this difference between the mean responses by the mean response across the control replicates.

$$0.095 \div 0.565 = 0.168$$

3. Multiply the result by 100 and express the product as a percent difference in response.

$$0.168 \times 100 = 16.8\% \text{ difference in response}$$

4. If the percent difference in response is≤ 39%, then the WET test has met the power standard.

A 16.8% difference in response is < 39%; the WET test has met the power standard.

Appendix F
Rapid Screening Tests and Species

1. Acute Rapid Screening Tests

Rapid screening tests for acute toxicity are expected to have a maximum mortality proportion of 0.20 in 100 percent effluent. The mortality proportion is calculated by subtracting the number of test organisms living in 100 percent effluent at the end of the test from the number of test organisms living in the control and dividing the result by the number of test organisms living in the control (Abbott's correction). The 100 percent effluent test concentration and the control must have equal numbers of test organisms.

A. Rotifer

The rotifer (*Brachionus sp.*) method is ASTM E 1440-91. The test is a 24-hr acute test using rotifers hatched from cysts. Tests with organisms hatched from cysts are less expensive because no time or materials are consumed by maintaining a culture. The rotifer test can be used in freshwater or saltwater.

B. 24-hour EPA Acute Screening Tests

The 24-hour EPA acute tests are conducted using the same EPA manual and species that were used for effluent characterization.

2. Chronic Rapid Screening Tests

A. Bacterial Bioluminescence Test (*Standard Methods* 8050)

B. Chronic Rotifer Test

The chronic rotifer test method is: Snell, Terry W. 1992. A 2-d Life Cycle Test With The Rotifer *Brachionus calyciflorus. Environ. Toxicol. Chem.* 11:1249-1257. The rotifer test measures the intrinsic rate of population increase. Measuring the intrinsic rate of population increase simultaneously evaluates both mortality and fecundity. Because it starts with rotifer cysts, uses small volumes of effluent, and only takes two days, it should be less expensive than EPA chronic tests.

C. Echinoderm Fertilization Test

The echinoderm fertilization rapid screening test method is: EPA/600/R-95/136. Because the fertilization test protocol is the same whether used for characterization, compliance monitoring, or as a rapid screening test, it is especially convenient.

Appendix G
Chlorine Toxicity

WET testing is not a good tool for regulating chlorine toxicity. The holding time for WET samples gives chlorine a longer time to react with organics or dechlorinating agents than occurs in the receiving environment. Chlorine is volatile. Steps taken to remove the supersaturation which occurs when cold samples are removed from storage and warmed to test temperature will also remove chlorine. Chlorine can be completely lost to aeration or reduced significantly as test solutions are poured into test chambers. Such a hit-or-miss situation is unfair to dischargers and labs who minimize holding times and sample handling and is not as protective as the water quality criteria for chlorine.

When chlorine is added to freshwater, the solution will contain two forms of free chlorine: hypochlorous acid (HOCl) and the hypochlorite ion (OCl^-). If the effluent also contains ammonia, then the addition of chlorine will result in two forms of combined chlorine: monochloramine and dichloramine. Municipal effluents usually contain all four of these forms of chlorine in some proportion and taken together they are known as "total residual chlorine" (TRC) and the EPA analytic method for TRC detects them in combination. Because saltwater contains bromide, the addition of chlorine to saltwater will also form hypobromous acid (HOBr), hypobromous ion (OBr^-), and bromamines. The term for the combination of chlorine and bromine compounds formed by the addition of chlorine to saltwater is "chlorine-produced oxidants" (CPO) and the EPA method for measuring total residual chlorine (TRC) also detects them.

The water quality criteria for chlorine in freshwater are based on total residual chlorine (TRC) and the criteria for saltwater are based on chlorine-produced oxidants (CPO). Both are measured, however, as total residual chlorine. The water quality criteria for chlorine in freshwater are: 19 µg/L (acute) and 11 µg/L (chronic). The criteria for saltwater are 13 µg/L (acute) and 7.5 µg/L (chronic). These criteria were calculated by U.S. EPA based on many toxicity tests on many species from both freshwater (33 animal species from 28 genera) and saltwater (24 animal species in 21 genera). Aquatic plants were less sensitive than aquatic animals and were not included in the calculations. Levels of TRC and CPO degrade very rapidly in water. In order to compensate for the degradation of TRC, CPO and their associated toxicity, U.S. EPA conducted the toxicity testing in the development of the water quality criteria for chlorine using flow-through systems with continuous introduction and monitoring of TRC during the test. The water quality criteria for chlorine are based on toxicity testing that is much more sensitive than the static or static-renewal tests used for effluent monitoring, and better protect surface waters from chlorine toxicity than the WET tests required in permits.

Other organochlorines formed by the chlorination of a complex effluent will not be detected by the method for total residual chlorine, but will also not affect WET. Scientists in the Environmental Investigations and Laboratory Services Program (EILS) of the Department of Ecology evaluated 16 POTW effluents sampled between February 1988 and August 1991 for 14 chlorinated organic compounds that were detected by chemical analysis. Only 4 of these chlorinated organic compounds appeared to be formed by effluent chlorination based on the observation that their concentrations were higher in the effluent than in the influent. These were chloromethane and three trihalomethanes (bromodichloromethane, dibromochloromethane, and chloroform). The 4 chlorinated organics presumed to be formed by effluent chlorination were orders of magnitude below water quality criteria for aquatic

life protection in every sample. These chlorinated organics in POTW effluent that are not detected when TRC is measured are also very unlikely to contribute to WET.

40 CFR 122.44(d)(1)(v) allows us to use chlorine limits instead of WET testing to regulate chlorine toxicity because our state has narrative water quality criteria for toxicity. To avoid the hit-or-miss detection of chlorine toxicity by WET testing and to avoid encouraging excessive use of dechlorinating agents by POTWs which control chlorine well enough to meet water quality standards at the edge of a mixing zone, we prefer that samples for WET testing be taken before the chlorinator for chlorinated discharges which can meet water quality-based effluent limits for chlorine and have an ACEC below 25% effluent. If a permit requires dechlorination of samples or if a permit requires sampling prior to the chlorinator and this is physically impossible, then the sample should be dechlorinated using a stoichiometrically determined amount of sodium thiosulfate or sulfur dioxide. The calculations for determining the amount of dechlorinating agent must be included in the test report. Because of the effluent-dominated receiving water condition when the ACEC is 25% effluent or higher, it is likely that permits will encourage extra control on chlorine through WET testingof an unmodified sample of final effluent.