Exhibit K

United States
Environmental Protection
Agency

Office of Water
(4303)

EPA 821-B-00-004
July 2000

EPA

# Method Guidance and Recommendations for Whole Effluent Toxicity (WET) Testing (40 CFR Part 136)

# 4 Concentration-Response Relationships

This chapter is designed to explain the concept of a concentration-response relationship. This chapter also identifies common patterns of WET test data and provides guidance on using the concentration-response concept to review WET test results.

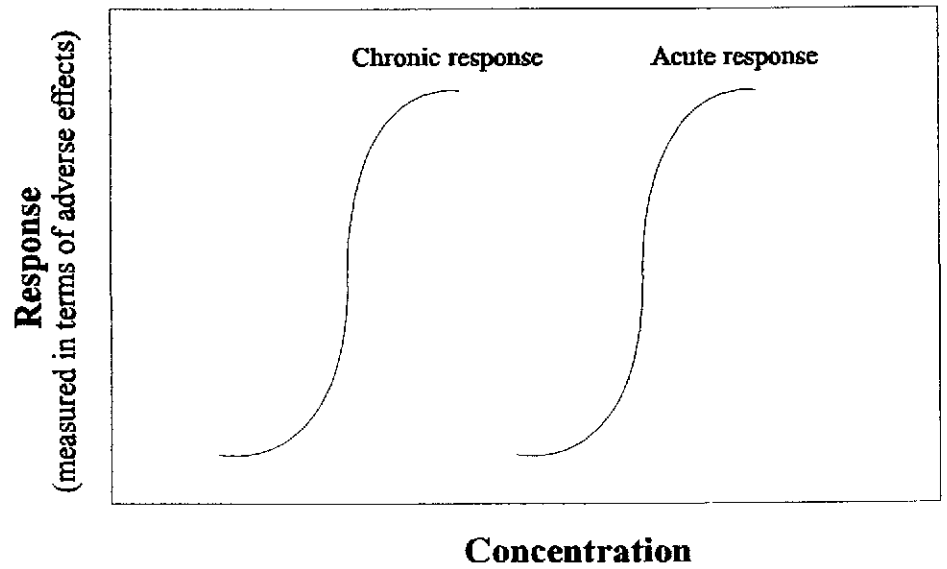
### How will this guidance be incorporated into WET test methodology?

EPA plans to incorporate the guidance presented in this chapter into the WET method manuals (USEPA, 1993c; USEPA, 1994a; USEPA, 1994b). A proposal to amend the manuals is expected to appear in the *Federal Register* by March 2001.

### What is the concentration-response relationship concept?

The concept of a concentration-response, or more classically, a dose-response relationship is "the most fundamental and pervasive one in toxicology" (Casarett and Doull, 1975). This concept assumes that there is a causal relationship between the dose of a toxicant (or concentration for toxicants in solution) and a measured response. A response may be any measurable biochemical or biological parameter that is correlated with exposure to the toxicant. The classical concentration-response relationship is depicted as a sigmoidal shaped curve (Figure 4.1), however, the particular shape of the concentration-response curve may differ for each coupled toxicant and response pair.

**Figure 4.1. Classical concentration-response relationship.**



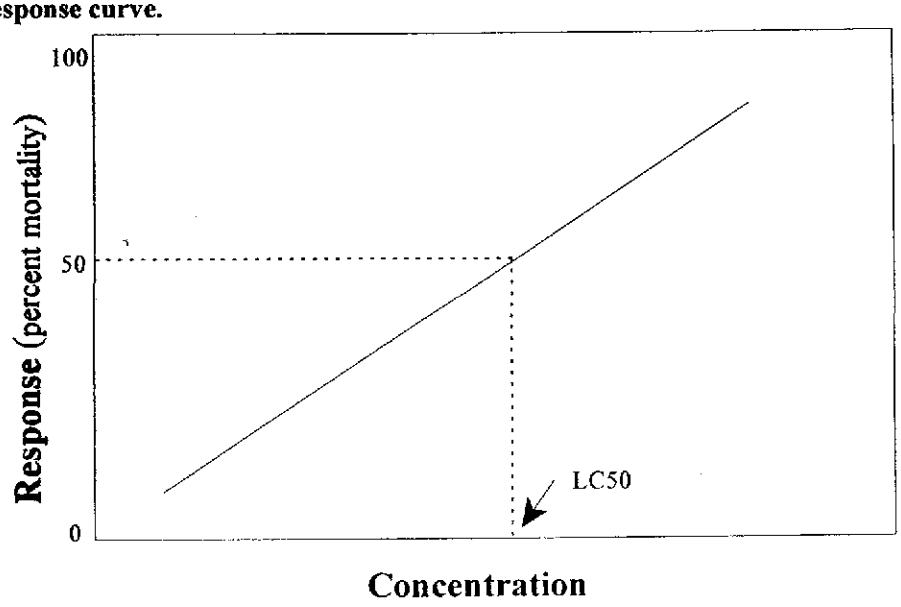
In general, more severe responses (such as acute effects) occur at higher concentrations of the toxicant, and less severe responses (such as chronic effects) occur at lower concentrations (Figure 4.1). A single toxicant also may produce multiple responses, each characterized by a concentration-response relationship.

In classical toxicology, concentration-response curves are generally displayed such that responses increase with increasing concentration (Figure 4.1). This is accomplished by defining responses in terms of adverse effects (e.g., mortality, reduction in growth, reduction in reproduction). The WET method manuals do not follow this convention; rather, responses are displayed in terms of survival, growth, and reproduction such that concentration-response curves for toxicants decrease with increasing concentration. This guidance will remain consistent with the convention established in the WET method manuals and will display concentration-response relationships for WET data such that responses decrease with increasing concentration.

### How is the concentration-response concept used in WET testing?

The concentration-response concept is the basis for the determination of point estimates (LC50, EC50, IC25, etc.) in WET testing. A biological response (mortality, growth inhibition, reproductive inhibition, etc.) is measured at a range of effluent concentrations to develop a concentration-response curve. This curve, which is typically sigmoidal, is then linearized by various transformations of the data (e.g., probit transform) to assist in drawing conclusions from the relationship. From the resulting linearized concentration-response curve, a point estimate effect concentration can be calculated (Figure 4.2). The effect concentration is an estimate of the concentration of effluent that will produce a specific level of response (e.g., 50% mortality). In WET testing, effect concentrations such as the LC50, EC50, IC25 and IC50 are commonly used to report WET test results.

**Figure 4.2. Example determination of point estimates from a concentration-response curve.**



## How can the concentration-response concept be used to review WET test results?

A corollary of the concentration-response concept is that every toxicant should exhibit a concentration-response relationship, given that the appropriate response is measured and given that the concentration range evaluated is appropriate. Use of this concept can be helpful in determining whether an effluent possesses toxicity and in identifying anomalous test results. An evaluation of the concentration-response relationship generated for each sample is an important part of the data review process that should not be overlooked. This chapter provides guidance on identifying valid concentration-response relationships and interpreting results from unexpected concentration-response patterns. This guidance on reviewing concentration-response relationships should be viewed as a component of a broader quality assurance and data review and reporting process that includes:

- **Review of test conditions** - The WET method manuals provide a summarized method-specific list of test conditions that should be followed in all WET test (e.g., test temperatures, number of replicates, test chamber sizes and volumes, lighting, feeding regimes, etc.). The conduct of each test should be reviewed to ensure that these conditions were met within the flexibility provided by the method manuals. The test conditions used in the test and any deviation from WET method manual requirements should be clearly reported. Daily measurements should be reviewed to ensure that values are within the acceptable ranges. Calibration of equipment should be verified and noted.
- **Review of test acceptability criteria** - The WET method manuals provide method-specific minimum criteria for the acceptability of tests (e.g., minimum control survival, reproduction, growth, or variability). These criteria are requirements of the methods, and any test not meeting the minimum test acceptability criteria should be considered invalid. All invalid tests should be repeated with a newly collected sample. While permit compliance should not be based on an invalid test, EPA's promulgation of the methods requires the results of all tests to be reported (valid or invalid).
- **Review of reference toxicant testing** - Reference toxicant testing is an important quality control practice that is required in the WET method manuals. Reference toxicant testing should be conducted on at least a monthly basis for each test method routinely conducted in a laboratory. WET test review should include evaluation of the most recent reference toxicant test and the reference toxicant cusum chart maintained by the laboratory. All reference toxicant tests should be conducted similarly (e.g., test duration, test conditions, test endpoint) to effluent tests being conducted. For instance, acute reference toxicant testing should be conducted to accompany acute testing of effluents, and short-term chronic reference toxicant testing should be conducted to accompany short-term chronic testing of effluents.
- **Review of organism culture health and performance** - EPA recommends that laboratories monitor and record the health and performance of organism cultures from which test organisms are obtained. For instance, the survival and reproduction of *Ceriodaphnia dubia* brood stock should be monitored and recorded during routine culture maintenance (i.e., water changes). This can be accomplished with a subset of 10 to 20 brood culture animals in individual culture vessels. This monitoring and documentation allows a laboratory to assess the current condition

of organism cultures prior to initiating a test and can allow the laboratory to postpone testing if organism cultures are unhealthy. This can potentially reduce the incidence of invalid tests and the cost associated with retesting. In the test review step, the documentation of culture health and performance can be useful in either identifying or eliminating poor culture health as a cause for marginal control performance in a test. Laboratories should maintain culture control charts (cusum charts) for survival, reproduction, growth, or other parameters for the appropriate species.

- **Review of test variability** - EPA recommends that the variability of each WET test, measured as a minimum significant difference (MSD) or percent MSD, be calculated and reported with all test results. EPA also recommends that laboratories maintain control charts for percent MSDs (USEPA, 2000). These control charts will allow laboratories to assess individual test variability in the context of typical variability within the laboratory. High test variability can result in insensitive tests or unexpected concentration-response relationships. Consult USEPA (2000) for additional guidance on WET test method variability.
- **Review of concentration-response relationships** - The guidance provided in this chapter may be used to assist in evaluating the concentration-response relationship as a part of the data review and reporting process. The succeeding section ("What are some patterns of concentration-response relationships typically seen in WET test data?") provides examples of common patterns in WET test data, discusses possible causes and solutions for unexpected patterns, and provides guidance on when to accept or reject test data based on the concentration-response concept. Some states have already developed similar guidance (Washington State Department of Ecology, 1997). It should be noted that the determination of a valid concentration-response relationship is not always clear cut. Data from some tests may suggest consultation with professional toxicologists and/or regulatory officials. Tests that exhibit unexpected concentration-response relationships also may indicate a need for further investigation and possible retesting. In general, when unexpected or apparently anomalous concentration-response relationships are encountered, EPA recommends the following:
  - **attempt to determine a cause for the response** - The above mentioned test review steps and specific guidance for individual concentration-response relationships (see "What are some patterns of concentration-response relationships typically seen in WET test data?") may assist in determining a cause for unexpected concentration-response relationships. Unexpected concentration-response relationships could be valid response patterns or anomalies resulting from Type I test error, high test variability, or other causes. If a given effluent consistently produces a specific, unexpected concentration-response relationship, there is likely a physical, chemical or biological cause. In situations where difficult-to-interpret concentration-response relationships are produced consistently by a given effluent, consultation with professional toxicologists is recommended. Toxicity identification evaluation (TIE) procedures (USEPA, 1991a; USEPA, 1992; USEPA, 1993a; USEPA, 1993b; USEPA, 1996b) also provide guidance that may be useful in determining a cause for such concentration-response relationships.
  - **follow guidance for specific concentration-response patterns** - The succeeding section ("What are some patterns of concentration-response relationships typically seen in WET

test data?") provides examples of 10 concentration-response patterns that may be exhibited by WET test data. This section provides guidance in interpreting each concentration-response pattern using a step-by-step review process. Based on this review, the guidance may recommend acceptance of the calculated results (e.g., NOEC or IC25) as valid and reliable, explanation of the calculated results as anomalous, or retesting.

- **increase testing frequency** - EPA recommends a testing frequency increase after any anomalous, questionable, or failing test result, with the number of tests and duration of testing to be determined by the regulatory authority.
- **coordinate with regulatory authorities, permittees, and testing laboratory** - EPA recommends that regulatory authorities, permittees, and testing laboratory personnel work together to resolve difficult-to-interpret WET test data. EPA also recommends that discussions be initiated as soon as possible when questions arise regarding WET test results.

This chapter provides additional guidance on reviewing test data; it is not the intent of this chapter to recommend the frequent disqualification and repetition of WET tests. Several warnings and safeguards should be considered when implementing the guidance in this chapter. First, unexpected concentration-response relationships should not occur with any regular frequency. Second, it is not recommended to screen only those tests in which toxicity is found at or below the receiving water concentration (RWC). If screening is to be done for unexpected concentration-response relationships, all tests should be screened in a similar manner. Third, all testing results should be reported to the regulatory authority, and the regulatory authorities should review all tests (including those disqualified and repeated). Regulatory authorities should be alert to patterns such as a high or increasing test rejection rate or a tendency for disqualified tests to show toxicity more often than tests accepted without qualification.

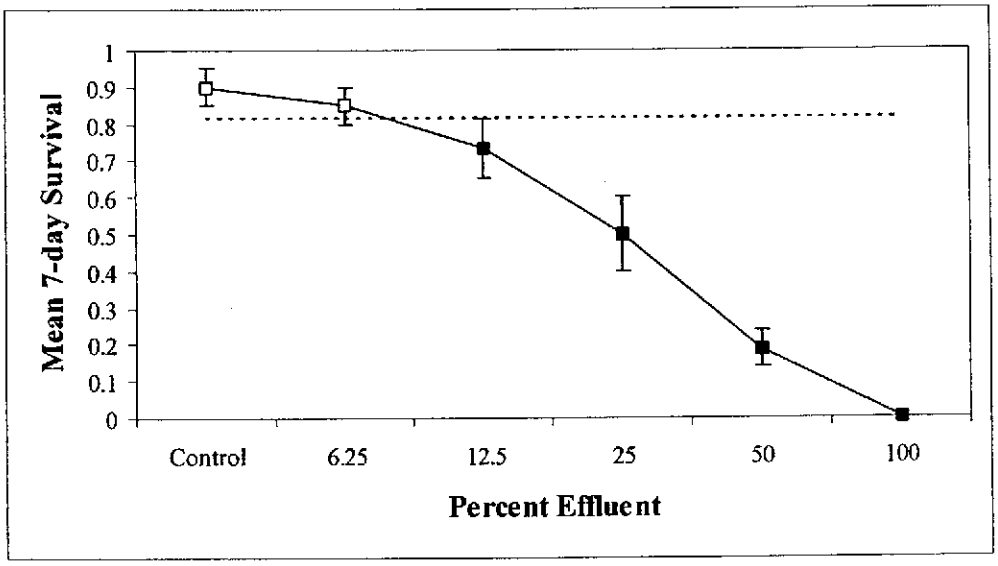
## What are some patterns of concentration-response relationships typically seen in WET test data?

Ten concentration-response patterns that may appear in WET testing are individually described and illustrated below using hypothetical test data. This section provides guidance in interpreting each concentration-response pattern. The guidance focuses on determining a cause for unexpected concentration-response patterns by recommending a step-by-step review process. Based on this review, the guidance may recommend acceptance of the calculated results (e.g., NOEC or IC25) as valid and reliable, explanation of the calculated results as anomalous, or retesting. When retesting is recommended, this generally means beginning a new test on a newly collected sample since sample holding times are typically expired by the time results are obtained from the original test. Test results should be reported for all tests conducted, even if retesting is recommended.

### 1. Ideal concentration-response relationship

This response pattern (Figure 4.3) shows a clear concentration-response relationship, with multiple effluent concentrations identified as significantly different from the control. This pattern also shows a monotonic decrease in response, meaning that the response steadily decreases for each higher effluent concentration. This pattern is indicative of a well designed test with appropriately chosen concentrations that bracket the effluent's range of toxicity. Under these circumstances, the hypothesis testing and point estimation techniques recommended in the WET method manuals provide reliable results.

**Figure 4.3. Ideal concentration-response relationship.** [1]

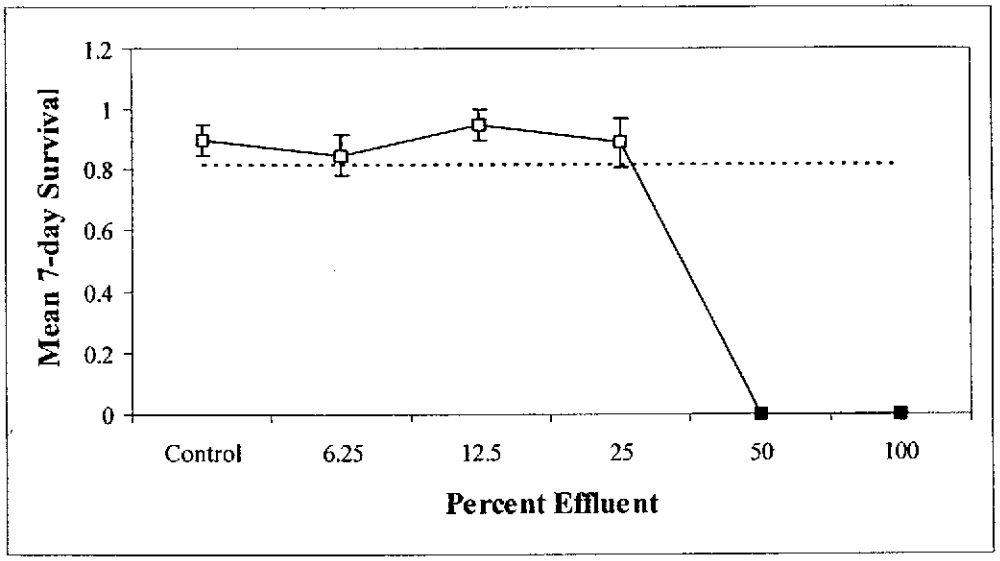



[1] Solid squares indicate data points that are statistically significantly different from the control, and hollow squares indicate data points that were not significantly different from the control. The dotted line shows the control mean minus the minimum significant difference (MSD); any test treatment response mean less than this value is considered to differ significantly from the control mean.

### 2. All or nothing response

The "all or nothing" response pattern is very common in WET test data. This response pattern (Figure 4.4) is characterized by a transition from no significant effect at one effluent concentration to a complete effect (100% mortality) at the next higher concentration. While not ideal, this pattern also represents a valid concentration-response relationship, and both hypothesis testing and point estimation techniques recommended in the WET method manuals will provide reliable results. This pattern of response is indicative of a steep concentration-response curve for the given effluent, and under these circumstances, the precision of the estimate may be improved by closer spacing of effluent concentrations (increased dilution factor) or the addition of intermediate effluent concentrations in future testing.

**Figure 4.4. All or nothing concentration-response relationship.** [1]

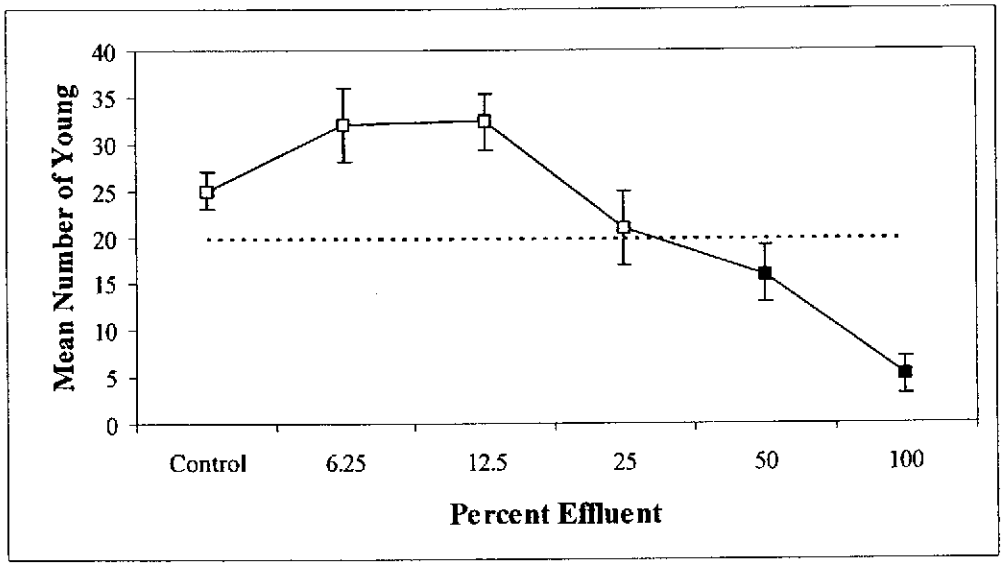



[1] Solid squares indicate data points that are statistically significantly different from the control, and hollow squares indicate data points that were not significantly different from the control. The dotted line shows the control mean minus the minimum significant difference (MSD); any test treatment response mean less than this value is considered to differ significantly from the control mean.

### 3. Stimulatory response at low concentrations and detrimental effects at higher concentrations

A stimulatory response is a nonmonotonic concentration-response relationship characterized by a measured increase in the response (stimulation) at low concentrations. This stimulation at low concentrations can be followed by a detrimental effect at higher concentrations (Figure 4.5) or by no effect at higher concentrations (see Section 4 following). Davis and Svendsgaard (1993) found that such nonmonotonic concentration-response relationships occurred in 12-24% of the toxicological studies surveyed. The stimulatory response pattern characterized in Figure 4.5 is typically found with sublethal endpoints such as reproduction, growth, fertilization, or larval development. For instance, test organism reproduction may increase (relative to the control) at low concentrations of an effluent and decrease relative to the control at higher concentrations. This concentration-response pattern, while nonmonotonic, is still a valid concentration-response relationship, and both hypothesis testing and point estimation techniques recommended in the WET method manuals will provide reliable results.

**Figure 4.5. Stimulation at low concentrations and significant effects at high concentrations.** [1]

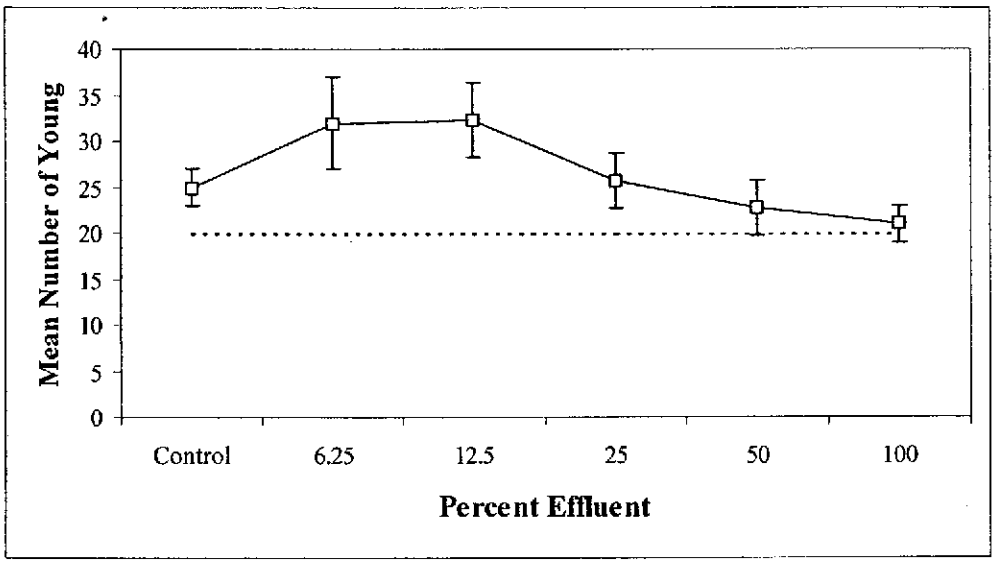



[1] Solid squares indicate data points that are statistically significantly different from the control, and hollow squares indicate data points that were not significantly different from the control. The dotted line shows the control mean minus the minimum significant difference (MSD); any test treatment response mean less than this value is considered to differ significantly from the control mean.

**4. Stimulation at low concentrations but no significant effect at higher concentrations**

This concentration-response relationship is similar to the previous example in that stimulation is observed at lower concentrations, but in this case, higher concentrations do not produce significant effects (Figure 4.6). In this situation, hypothesis testing techniques should produce reliable results, assuming that adequate test sensitivity is achieved. Results from point estimation techniques should be interpreted carefully when this response pattern is encountered, because the inhibition concentration percentage (ICp) procedure may produce effect concentrations (particularly IC25s) that indicate toxicity at effluent concentrations where the response is comparable to the control response. The ICp procedure assumes that responses: (1) are from a random, independent, and representative sample of test data; (2) follow a piecewise linear response function; and (3) are monotonically non-increasing, meaning that the mean response for each higher concentration is less than or equal to the mean response for the previous concentration. If the data are not monotonically non-increasing, the ICp procedure adjusts the response means using a "smoothing" technique that averages adjacent means (see Appendix M of USEPA, 1994a). This technique averages response means (including that of the control) with those of the next highest test concentration until responses are monotonically non-increasing. In cases where the responses at the low effluent concentrations are much higher than in the control, the smoothing process may result in a large upward adjustment in the control mean. This can lead to an IC25 result that is less than the highest test concentration, even though the highest test concentration was not statistically different from the control treatment and even if a percent difference of less than 25% was observed between the control response and the response at the highest test concentration.

**Figure 4.6. Stimulation at low concentrations but no significant effect at higher concentrations.** [1]




[1] Solid squares indicate data points that are statistically significantly different from the control, and hollow squares indicate data points that were not significantly different from the control. The dotted line shows the control mean minus the minimum significant difference (MSD); any test treatment response mean less than this value is considered to differ significantly from the control mean.

If the response pattern depicted in Figure 4.6 (stimulation at low concentrations but no significant effect at higher concentrations) is encountered, the following review steps should be taken in addition to standard test review procedures:

- **Evaluate the concentration range** - If the highest concentration used in the test was less than 100% effluent (or the highest achievable effluent concentration for marine tests), the effluent should be retested using higher test concentrations to establish if a valid concentration-response relationship exists. This may not be necessary if the permit limit is set at much lower than 100% effluent and test results indicate no toxicity at the permit limit level and at least one concentration above the permit limit.
- **Compare hypothesis testing results and point estimates** - If there is agreement between the NOEC and the IC25 for tests producing the concentration-response pattern depicted in Figure 4.6 (i.e., neither value indicates toxicity at or below the permitted RWC, or both values indicate toxicity at or below the RWC) the test results should be reported and considered valid. If, however, the NOEC indicates no toxicity at the RWC (i.e., NOEC greater than or equal to RWC) but the IC25 is calculated as less than the RWC, the remaining recommended actions should be taken.
- **Evaluate control response** - It is possible that the response pattern depicted in Figure 4.6 could result from poor performance in the controls rather than stimulation at the lower test concentrations. This poor control performance could cause a toxic effect at higher test concentrations not to be detected. To evaluate this possibility, compare the control response

to the normal control performance for the laboratory. If (1) a particular test exhibits the response pattern depicted in Figure 4.6, (2) there is disagreement between NOEC and IC25 estimates, and (3) the mean control response is well below the laboratory's normal range of control performance; retesting of the effluent is recommended even if the minimum test acceptability criteria have been met. For example, if a laboratory consistently achieves a control mean of 25-30 neonates for the *Ceriodaphnia dubia* 3-brood chronic test, a control mean of 15-18 neonates (in conjunction with a non-ideal concentration-response curve and disagreement between the NOEC and IC25) would warrant retesting. In this situation, suppressed control performance could be considered as the cause for this response pattern rather than stimulation. A review of control performance should also investigate the possibility of poor performance in a single replicate substantially reducing the mean control response. In this case, retesting is also recommended.

- **Evaluate the test sensitivity** - Discrepancies between IC25 and NOEC values could be due to low test sensitivity. To determine if this is the case, evaluate the sensitivity of the test by comparing the test MSD to MSD criteria for the given test method (see Chapter 2 of this guidance and USEPA, 2000) and to the laboratory's historical test sensitivity performance. Laboratories are encouraged to track test sensitivity (as %MSDs) for tests conducted over time. If a test exhibits the response pattern depicted in Figure 4.6 and the test MSD is above maximum recommended criteria for the method or above the laboratory's typical range, the sample should be retested.
- **Evaluate the ICp calculation** - If a test exhibits the response pattern depicted in Figure 4.6 and it has been determined from the above actions that the pattern is not due to poor control performance or low test sensitivity, then discrepancies between the NOEC and IC25 may be due to bias from the ICp smoothing technique. To determine if this is the case, calculate the observed percent difference between the response at the RWC and the control as:

$$\frac{\left(\mu_c - \mu_{RWC}\right)}{\mu_c} \times 100\%$$

where:

$\mu_c$ = mean control response

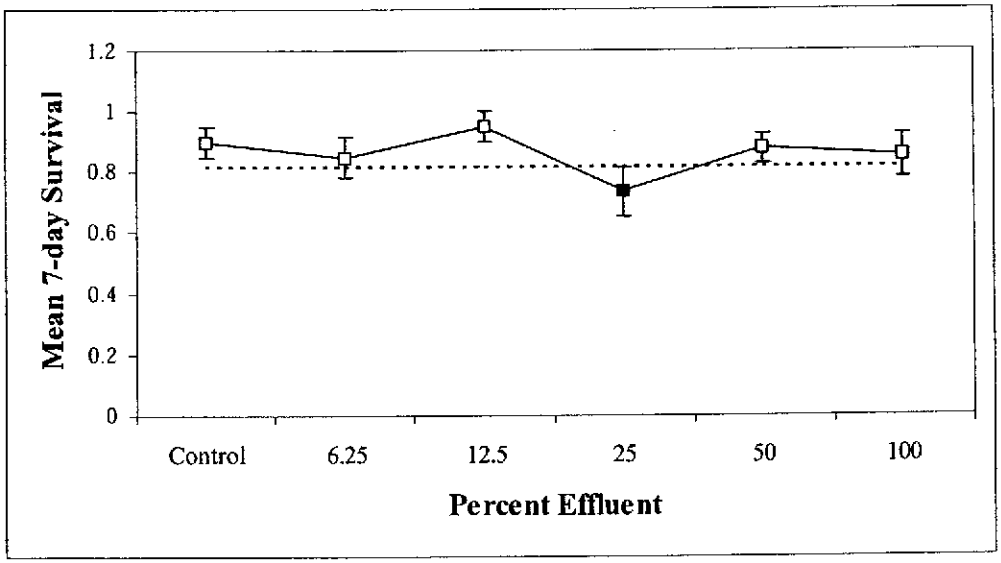
$\mu_{RWC}$ = mean response at the receiving water concentration (RWC)

If the observed percent difference between the response at the RWC and the control is less than 25% and the response at the RWC is not statistically significantly different from the control response, then a calculated IC25 of less than the RWC should be noted as anomalous and the effluent determined to be non-toxic at the RWC. If the observed percent difference is equal to or greater than 25%, then the calculated IC25 should be considered valid.

**5. Interrupted concentration-response: significant effect bracketed by non-significant effects**

This response pattern is characterized by a single test concentration showing a significant difference from the control while adjacent higher and lower test concentrations do not differ significantly from the control (Figure 4.7). When this response pattern is encountered, point estimation techniques generally will yield reliable results, but hypothesis testing results should be interpreted carefully. The method manual definitions of NOEC (the highest concentration of toxicant in which the values for the observed responses are not statistically significantly different from the controls) and LOEC (the lowest concentration of toxicant in which the values for the observed responses are statistically significantly different from the controls) were intended for situations where the concentration-response relationship is monotonically non-increasing. Under these circumstance, the NOEC and LOEC are always adjacent values with the NOEC being the test concentration just below the LOEC. In circumstances where the concentration-response relationship is non-monotonic (as in Figure 4.7), the identification of NOEC and LOEC values is severely compromised (Chapman *et al.*, 1996). For this response pattern, the following review actions should be taken in addition to standard test review procedures to determine the validity of results obtained by hypothesis testing:

**Figure 4.7. Interrupted concentration-response: significant effect bracketed by non-significant effects.** [1]




[1] Solid squares indicate data points that are statistically significantly different from the control, and hollow squares indicate data points that were not significantly different from the control. The dotted line shows the control mean minus the minimum significant difference (MSD); any test treatment response mean less than this value is considered to differ significantly from the control mean.

- **Check for test condition or procedural errors** - The concentration-response relationship depicted in Figure 4.7 could result from test conditions errors (such as pH, DO, salinity, or temperature excursions) occurring in isolated test replicates. This concentration-response

pattern also could be due to procedural errors such as failure to properly randomize test organisms or test chamber placement. The laboratory should verify that all test conditions were within ranges required by the WET method manuals for the given test method. The laboratory should verify that the assignment of test organisms to individual treatments was properly randomized (Davis *et al.*, 1998). This can be complete randomization or block randomization (as with the *Ceriodaphnia dubia* 3-brood reproduction test). The laboratory also should verify that the positions of test chambers within the experiment were properly randomized. If test condition or procedural errors are identified, the sample should be retested.

- **Evaluate within-treatment variability** - It is possible for poor performance in a single replicate to bias the mean response for a given test concentration and cause that concentration to differ significantly from the control. For this reason, the within-treatment variability should be evaluated for the significantly different treatment. If the variability (standard deviation or CV) for that treatment is considerably greater than for other treatments, then responses of individual replicates should be investigated. This investigation may show that a single outlier replicate has biased the treatment mean. If this is the case and the responses from all but the single outlier replicate are consistent with the control response, then the sample should be retested.
- **Evaluate test sensitivity** - When the response pattern depicted in Figure 4.7 is encountered, it is important to evaluate test sensitivity. If test sensitivity is low (e.g. high MSD values), large effects at higher test concentrations may not be detected as statistically significant. To evaluate test sensitivity, compare the MSD for the test to benchmark criteria for the given test method (see Chapter 2 of this guidance and USEPA, 2000) and to the laboratory's historical test sensitivity performance. As previously mentioned, laboratories are encouraged to track test sensitivity (as %MSDs) for tests conducted over time. If test sensitivity is low (i.e., MSDs are above maximum recommended criteria or typical laboratory performance), then the sample should be retested. Consult Section 6.4 in USEPA (2000) for additional guidance on implementing upper and lower bounds on test sensitivity.

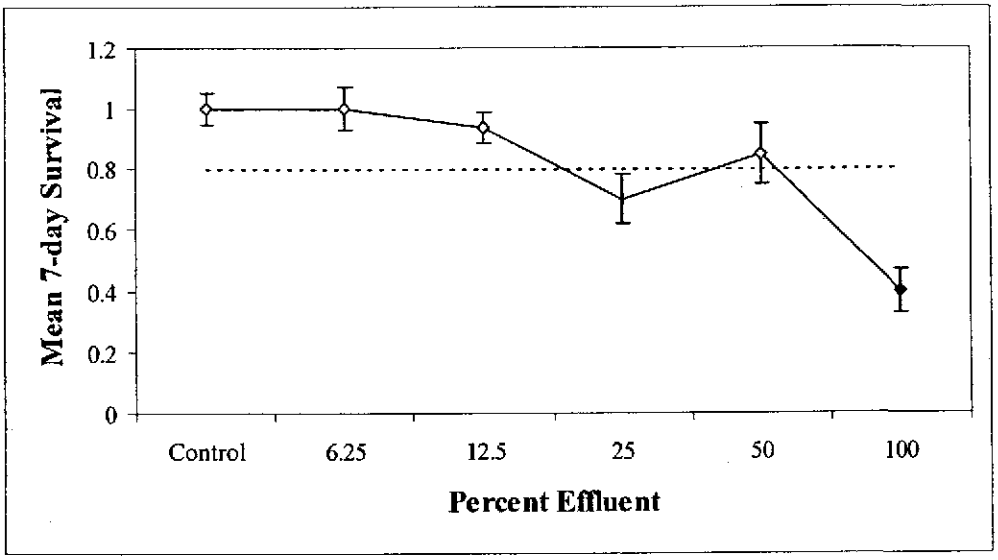
If test sensitivity is moderate to high (i.e., MSDs below the maximum recommended criteria and within the laboratory's typical performance range) and none of the preceding evaluations have determined a cause for this response pattern, it is likely that the significantly different treatment is the result of a Type I error. A Type I error is the error of incorrectly rejecting the null hypothesis (assuming that the treatment is significantly different from the control) when in fact the null hypothesis is true (the treatment is not significantly different from the control). In this situation, due to the absence of a valid concentration-response relationship, the intermediate concentration that was determined by hypothesis testing to be statistically different from the control should be considered anomalous, and the NOEC should be determined as the highest concentration that was not significantly different from the control. Using Figure 4.7 to illustrate, the 25% concentration would be considered anomalous, the reported NOEC would be 100%, and the reported LOEC would be >100%. Under these circumstances, test results should still note that the 25% concentration was statistically

different from the control but was considered anomalous due to analysis of the concentration-response curve and the above review steps.

**6. Interrupted concentration-response: non-significant effects bracketed by significant effects**

This response pattern is similar to the previous response pattern in that the concentration-response curve is nonmonotonic (or interrupted), however, this response pattern is characterized by two or more test concentrations showing a significant difference from the control while an intermediate test concentration does not differ significantly from the control (Figure 4.8). When this response pattern is encountered, point estimation techniques will generally yield reliable results, but hypothesis testing results should be interpreted carefully. As mentioned for the previous concentration-response pattern, the identification of NOEC and LOEC values is severely compromised (Chapman *et al.*, 1996) when the concentration-response relationship is non-monotonic (as in Figure 4.8). For this response pattern, the test sensitivity should be evaluated as described below in addition to standard test review procedures to determine the validity of results determined by hypothesis testing.

**Figure 4.8. Interrupted concentration-response: non-significant effects bracketed by significant effects.** [1]




[1] Solid squares indicate data points that are statistically significantly different from the control, and hollow squares indicate data points that were not significantly different from the control. The dotted line shows the control mean minus the minimum significant difference (MSD); any test treatment response mean less than this value is considered to differ significantly from the control mean.

- **Evaluate test sensitivity** - When the response pattern depicted in Figure 4.8 is encountered, it is important to evaluate test sensitivity by comparing test MSDs to minimum and maximum MSD criteria recommended by EPA (USEPA, 2000). If the test MSD is lower than the minimum MSD criterion, only effects larger than the minimum MSD criterion should be

considered significant. For example, if the minimum MSD criterion for a method is 15% and the calculated test MSD is 10%, only effects greater than 15% difference compared to the control should be considered significant. If test sensitivity is low (i.e., test MSD is above maximum MSD criterion), the sample should be retested. If test sensitivity is moderate (i.e., test MSD is within minimum and maximum MSD criterion), the test results should be considered valid and the NOEC should be reported as the concentration below the LOEC. For the case depicted in Figure 4.8, a NOEC of 12.5% should be reported. Consult Section 6.4 in USEPA (2000) for additional guidance on implementing upper and lower bounds on test sensitivity.

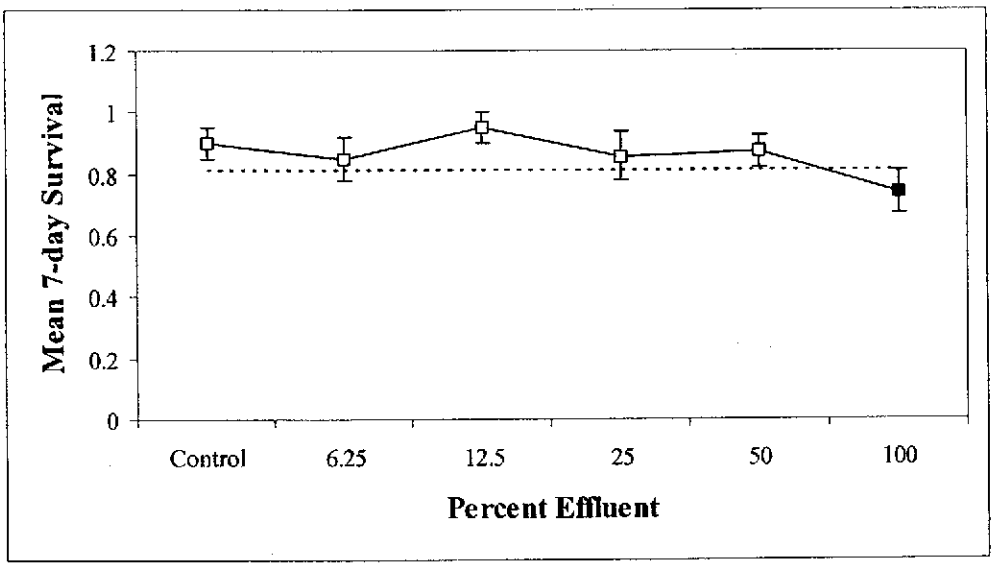
**7. Significant effects only at highest concentration**

This response pattern is characterized by only the highest test concentration producing a significantly different response from the control (Figure 4.9). This response pattern should be considered to be a valid concentration-response relationship and results determined by point estimation should be assumed to be reliable. Hypothesis testing results are also assumed to be reliable following the evaluation of test sensitivity as described below. If the response pattern depicted in Figure 4.9 (significant effects only at highest concentration) is encountered, the following review steps should be taken in addition to standard test review procedures:

- **Evaluate the concentration range** - When this response pattern occurs, the concentrations used for testing should be evaluated in future tests using this effluent. If the highest effluent concentration used in the test was less than 100% (or the highest achievable effluent concentration for marine tests), future testing using this sample should include at least one higher test concentration to confirm the presence of a concentration-response relationship. If the test used a 100% effluent concentration treatment, it is difficult to confirm a concentration-response relationship through retesting because concentrations are constrained to less than or equal to 100% in whole effluent testing. If this response pattern occurs commonly with a given effluent, future testing of the effluent should use a dilution factor of >0.5 such that test concentrations closer to the 100% effluent concentration are used (i.e., a dilution factor of 0.65 would provide a test concentration series of 18%, 27%, 42%, 65%, and 100%). This would provide a better opportunity to confirm a concentration-response relationship that may exist at the upper end of the concentration range. This approach should be used only if historical testing of the effluent indicates consistency and the effect concentration is not likely to fall below the adjusted test concentration series.
- **Evaluate test sensitivity** - Evaluate test sensitivity by comparing test MSDs to minimum and maximum MSD criteria recommended by EPA (USEPA, 2000). If the test MSD is lower than the minimum MSD criterion, only effects larger than the minimum MSD criterion should be considered significant. For example, if the minimum MSD criterion for a method is 15% and the calculated test MSD is 10%, only effects greater than 15% difference compared to the control should be considered significant. If test sensitivity is low (i.e., test MSD is above maximum MSD criterion), the sample should be retested. If test sensitivity is moderate (i.e., test MSD is within minimum and maximum MSD criterion), the test results should be

considered valid and the NOEC should be reported as the concentration below the LOEC. For the example given in Figure 4.9, a NOEC of 50% effluent should be reported. Consult Section 6.4 in USEPA (2000) for additional guidance on implementing upper and lower bounds on test sensitivity.

**Figure 4.9. Significant effects only at highest concentration.** [1]




[1] Solid squares indicate data points that are statistically significantly different from the control, and hollow squares indicate data points that were not significantly different from the control. The dotted line shows the control mean minus the minimum significant difference (MSD); any test treatment response mean less than this value is considered to differ significantly from the control mean.

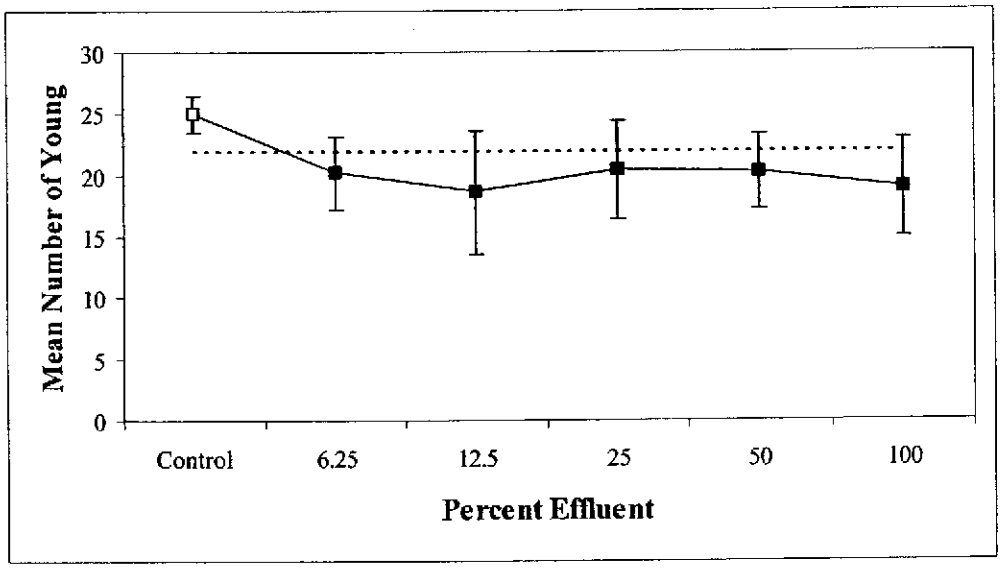
**8. Significant effects at all test concentrations but flat concentration-response curve**

This response pattern is demonstrated in Figure 4.10. All of the test concentrations produce a response that is significantly different from the control response, but a clear concentration-response relationship cannot be determined. This response pattern could be due to: (1) extremely low variability in the control, (2) an unusually high control response, (3) an inappropriate dilution water and improper use of dilution water controls, (4) inappropriate test dilution series, (5) potential pathogen effects in the effluent, (6) an unusual effluent-dilution water interaction. The following review actions should be taken to determine a cause for this concentration-response pattern and to subsequently determine the validity of calculated results.

- **Evaluate test sensitivity** - The response pattern depicted in Figure 4.10 may be an artifact of the data resulting from extremely precise control results and extremely high test sensitivity. Investigate this possibility by comparing test MSDs to minimum MSD criteria recommended by EPA (USEPA, 2000). If the test MSD is lower than the minimum MSD criterion, only effects larger than the minimum MSD criterion should be considered significant. For example, if the minimum MSD criterion for a method is 15% and the calculated test MSD is 10%, only effects greater than 15% difference compared to the control should be considered significant.

If test sensitivity is low (i.e., test MSD is above maximum MSD criterion), the sample should be retested. Consult Section 6.4 in USEPA (2000) for additional guidance on implementing upper and lower bounds on test sensitivity.

**Figure 4.10. Significant effects at all test concentrations but flat concentration-response curve.** [1]




[1] Solid squares indicate data points that are statistically significantly different from the control, and hollow squares indicate data points that were not significantly different from the control. The dotted line shows the control mean minus the minimum significant difference (MSD); any test treatment response mean less than this value is considered to differ significantly from the control mean.

- **Evaluate control response** - The concentration-response pattern depicted in Figure 4.10 could result from an unusually high response in the control treatment. Laboratories are encouraged to track the performance of controls in tests conducted over time. When the response pattern depicted in Figure 4.10 is exhibited, the control response for the test should be compared to historic control performance in the laboratory using the given dilution water. If the mean control response is above the normal range for that laboratory and dilution water, the sample should be retested.
- **Evaluate dilution water** - The improper use of dilution waters and dilution water controls could cause the concentration-response pattern depicted in Figure 4.10. It should be confirmed that test treatment concentrations were compared to the dilution water control and not a culture water control. A statistical comparison of the dilution water control and the culture water control should also be made if they are from different sources. If the dilution water control shows a statistically significant difference from the culture water control, alternate dilution waters should be considered and the sample retested (see Chapter 6 of this guidance).
- **Evaluate test concentrations** - If all test concentrations produce a complete effect (e.g., 100% mortality, zero reproduction, etc.), a flat concentration-response relationship will result. This concentration-response relationship should be considered valid, and it indicates high toxicity in

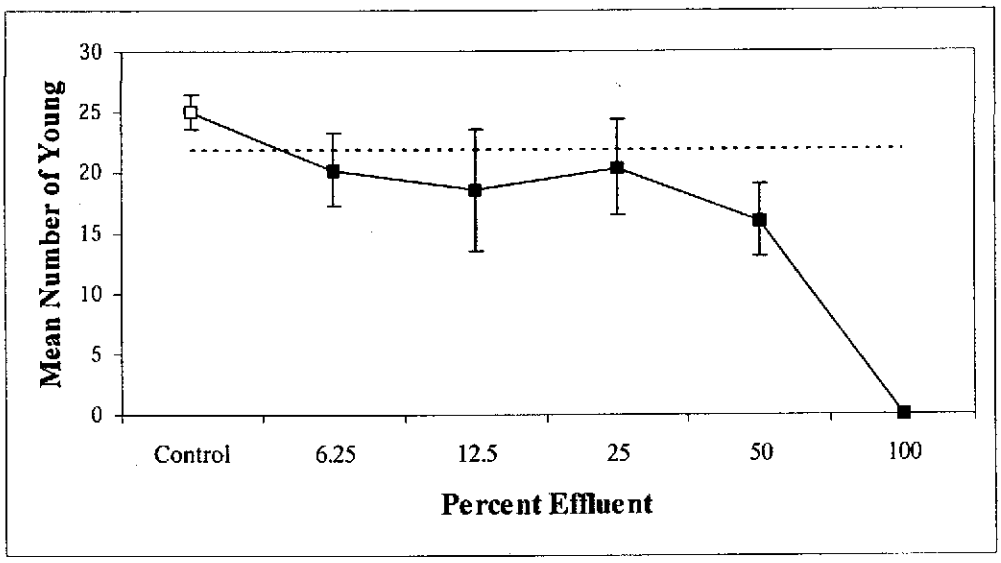
the sample. Assuming that the concentration range used in the test brackets the permitted RWC, it is not necessary to retest the sample, since the test results clearly indicate toxicity. If all test concentrations were significantly different from the control but did not produce complete effects (as in Figure 4.10), the dilution series should be investigated. It is possible that the test concentration range used for the test was too narrow to distinguish a shallow sloped concentration-response curve. Test concentrations may not have been low enough to produce no significant effect and may not have been high enough to produce severe effects. If this situation is suspected, the sample should be retested using an expanded dilution series range. Effluent concentrations that are lower than those used in the previous test should be added. Effluent concentrations that are higher than those used in the previous test also should be added (if possible) to assist in determining a concentration-response relationship.

- **Consider pathogen effect** - The concentration-response pattern depicted in Figure 4.10 could also be due to the presence of pathogens in the effluent. The most common identifier of pathogen effects are sporadic mortalities and extremely high variability between replicates. The pathogen effect is more common in tests using fish species than in invertebrate testing. This pathogen effect also may be evident only in chronic tests and not in acute tests. Pathogen effects also may be seasonal in occurrence. If within-treatment CVs for survival are >40% for effluent concentrations and relatively small for control replicates in standard synthetic water, pathogen effect should be considered. If pathogen effects are suspected in the effluent, this may be confirmed in subsequent side-by-side testing using the effluent and the effluent treated by brief exposure to UV light or the addition of antibiotics, or increasing the number of replicates and using less test organisms in each replicate. If pathogen effects in the effluent are confirmed, the sample should be retested and the regulatory authority should be consulted prior to changing testing procedures.
- **Continued testing** - If all of the above scenarios have been investigated and have not revealed the cause of the response pattern, the results should be considered valid; however, continued testing should be initiated in an effort to identify the cause of the response pattern. If an effluent consistently exhibits this response pattern, additional investigations could include chemical analysis or initiation of TIE procedures.

**9. Significant effects at all test concentrations with a sloped concentration-response curve**

This concentration-response pattern is similar to the pattern identified in item #8 above except a concentration-response curve can be identified at the higher effluent concentrations (Figure 4.11). This pattern is considered to be a valid concentration-response relationship, and point estimation techniques will generally yield reliable results. Results determined by hypothesis testing techniques should be interpreted carefully, and the cause for significantly different effects at low concentrations should be investigated as described for the response pattern described in item #8.

**Figure 4.11. Significant effects at all test concentrations with a sloped concentration-response curve.** [1]




[1] Solid squares indicate data points that are statistically significantly different from the control, and hollow squares indicate data points that were not significantly different from the control. The dotted line shows the control mean minus the minimum significant difference (MSD); any test treatment response mean less than this value is considered to differ significantly from the control mean.

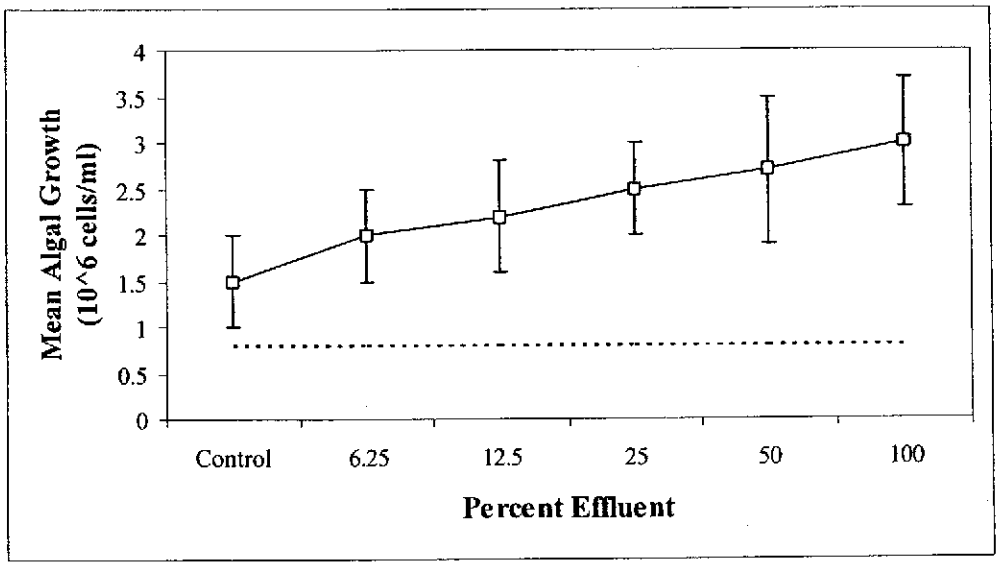
### 10. Inverse concentration-response relationship

This response pattern is characterized by a relationship in which adverse effects decrease with increasing effluent concentration (Figure 4.12). This situation is most often encountered in algal growth tests, and is typically caused by excess nutrients in the effluent. While a valid concentration-response relationship is demonstrated in this circumstance, the effluent should be considered nontoxic since the direction of the concentration-response relationship indicates decreasing adverse effects. It should be noted that while the effluent is considered non-toxic, the presence of excess nutrients still may pose a potential risk to the environment due to nutrient enrichment and oxygen depletion.

An inverse concentration-response pattern also may occur in tests other than algal growth assays when the dilution water used is a receiving water or synthetic water adjusted to approximate the receiving water characteristics. In such situations, the inverse concentration-response pattern can result from toxicity in the receiving water or the limitation of necessary components (i.e., hardness) in the receiving water or adjusted synthetic water. Under such circumstances, the objective of the toxicity test should be evaluated (see Chapter 6 of this guidance). If the objective of the test is to determine the toxicity of the effluent in the natural receiving water, then the results indicate no toxicity in the sample. If the objective of the toxicity test is to determine the absolute presence of toxicity in the effluent, the sample should be retested using a standard synthetic dilution water. Toxicity or limiting components in the receiving water or adjusted synthetic water may mask the

presence of low level toxicity in the effluent, making the absolute determination of toxicity in the effluent difficult.

**Figure 4.12. Inverse concentration-response relationship.** [1]




[1] Solid squares indicate data points that are statistically significantly different from the control, and hollow squares indicate data points that were not significantly different from the control. The dotted line shows the control mean minus the minimum significant difference (MSD); any test treatment response mean less than this value is considered to differ significantly from the control mean.

# 5 Dilution Series Selection

This chapter provides guidance on the selection of an appropriate dilution series for a WET test.

### Do the WET method manuals specify a certain dilution series?

The WET method manuals (USEPA, 1993c; USEPA, 1994a; USEPA, 1994b) suggest, but do not require, a dilution series of 6.25%, 12.5%, 25%, 50%, and 100% effluent for most effluents. This dilution series should be used as a default when little information is known about the effluent being tested and when initial range finding indicates that the effect concentration of interest is within the 6.25% to 100% effluent range. In many situations, a more appropriate dilution series can be selected based on experience from repeated testing of a given effluent. The WET method manuals do recommend a dilution factor of 0.5 for preparing test concentrations. This recommendation does not fix the dilution factor, but is provided to establish a lower limit on the dilution factor. The use of dilution factors greater than 0.5 is encouraged when historical testing indicates that an effluent is relatively consistent and effect concentrations generally fall within a given range.

### Why is selecting an appropriate dilution series important?

The selection of a dilution series (number and spacing of test concentrations) for WET tests is extremely important in producing reliable and precise results. This is most obvious for effect concentrations such as NOEC and LOEC values generated by hypothesis testing. These values are by definition limited to one of the effluent concentrations selected for the test. The precision of these values also is determined by the distance from the NOEC or LOEC to the next highest or lowest effluent concentration. For instance, using a standard dilution series of 6.25%, 12.5%, 25%, 50%, and 100% effluent, a measured NOEC value of 50% indicates that the transition from no observable effects to observable effects occurs somewhere between 50% and 100% effluent concentration (the NOEC-LOEC interval). If an alternative dilution series of 12.5%, 25%, 50%, 75%, and 100% were used for this test, then a NOEC of 50% would be a more precise estimate. In this test, the point of transition from no observable effect to observable effects is now known to lie between 50% and 75%.

The appropriate selection of a dilution series also is important for accurately identifying concentration-response relationships and increasing the precision of effect concentrations estimated from those relationships. For example, toxicants or effluents with steep concentration-response curves, often produce "all or nothing" results when using a standard dilution series of 6.25%,