May 30, 2003

Changes to the WET Test Methods

1. Specify a detection limit for each WET test and each endpoint that protects against false positives due to inherent variability in organism response.

Having a detection limit would be analogous to the MDL for traditional chemical methods. The current lack of detection limits for WET is, in our view, a critical flaw in the effort to transform WET from a tool for investigative work into enforceable NPDES permit conditions. Therefore, detection limits are critical to being able to utilize WET data fully.

2. Abandon hypothesis testing as a stand-alone or primary determinant of WET.

a. EPA might retain hypothesis testing, but only to confirm results obtained by point-estimation or direct measurement of biological effect.

b. Results from hypothesis testing must never be used by themselves for "reasonable potential" determinations, limit derivation, or numeric limit compliance.

3. Develop parametric point estimation models that work for all test endpoints and generate reliable confidence intervals.

a. None of EPA's models for continuous data is parametric. Some methods, such as the IC_p, assume that there is no fixed mathematical relationship among testing concentrations but require a monotonic relationship between concentrations. This requirement forces data into a model that cannot reliably represent them. EPA should provide parametric alternatives such as general linearized models.

b. In 1995 EPA recommended using point estimates (rather than NOECs) because confidence intervals can be placed around a point estimate (see 60 Fed. Reg. 53,539). In the final 2002 WET methods, however, EPA's method of calculating point estimates often fails to generate appropriate confidence intervals. EPA accepts this failure on the basis that confidence intervals are not reported in the Permit Compliance System or used to determine compliance. Confidence intervals should be used to account for the test result uncertainty when determining compliance.

4. Require a dose-response relationship.

a. EPA must include a statistically-based procedure in the methods to define this relationship and require the presence of this relationship to identify valid tests when toxicity is indicated in at least one dilution.

Chapter 4 of EPA's guidance document, *Method Guidance and Recommendations for Whole Effluent Toxicity (WET) Testing* (EPA 821-B-00-004 July 2000), describes how to evaluate ten data patterns. This would be the first step to be used with all data patterns. For example, WET data with a series of dilutions could be subjected to a linear regression analysis. A dose-response relationship would be inferred only if the slope of the regression line is negative and significantly different from zero at the 95% probability level. Retesting when a WET test does not meet this criterion must be required to determine the presence of toxicity.

b. Some test results should be declared anomalous.

The evaluation of some of the data patterns in Chapter 4 of the *Method Guidance* should be reevaluated and the reporting conditions changed. For example, we have specific technical problems with the fifth and sixth patterns in Chapter 4.

5. Withdraw the Federal Register language recommending the "West Coast methods" for limited, localized, or regional use (67 Fed. Reg. 69,955 col. 1-2 (Nov. 19, 2002)).

EPA did not approve the "West Coast methods" (e.g., *Holmesimysis costata*) because it did not have a minimum of six laboratories qualified and willing to perform the tests as part of the Interlaboratory Validation Studies. EPA should not support the use of test methods that have not been validated or approved for inclusion in Part 136.

6. Restore the method for calculating growth endpoints that was proposed in 1989.

a. In the 1995 version of the chronic methods, EPA adopted a procedure that was different from the procedure proposed for comment in 1989. The new procedure calculates growth based on the number of organisms starting a test instead of those surviving the test as in the 1989 proposal.

b. There is nothing in the record to show that sublethal endpoints need to be more sensitive. There was no comment from the public or scientific community urging EPA to make this change.

7. Add Data Quality Objectives (DQOs) to the WET methods.

Data Quality Objectives (DQOs), particularly "acceptance criteria," are necessary to determine whether WET monitoring data are suitable for their intended purpose. (This principle is discussed by EPA in its Guidance for the Data Quality Objectives Process, EPA QA/G-4 (August 2000), and in EPA Order 5360.1-A2, which requires EPA to establish DQOs.) For WET test method results to be sufficiently reliable for regulatory use, they must satisfy several DQOs, including mandatory testing protocols. EPA should do the following:

a. Require the use of WET Data Acceptance Criteria which address the same QA/QC issues raised in the attached checklist.

b. Identify national norms for all QA/QC metrics and establish acceptable ranges which must be met to validate sample results.

c. Clarify that inconsistent results from split samples are not a violation but instead may trigger additional testing.

d. Explain how to interpret results when stress and test interferences (pH shift, pH shock, ionic imbalance, or pathogens) are suspected, particularly when the problem cannot be eliminated entirely.

e. Revise the test acceptance criteria to account for natural sources of biological stress.

f. Provide upper and lower limits for the response of *controls* (nontoxic water).

   i. This would preclude unrepresentative organisms from influencing test results.

   ii. The upper and lower limits must be identified from control charts kept by each laboratory.

   iii. Keeping such charts and identifying upper and lower limits must be a mandatory QC requirement.

   iv. The upper and lower limits will be defined by a 95% confidence interval.

g. A reference toxicant which demonstrates that the control population of organisms is responding according to historical testing.

h. The IC_{25} can only be used with the current statistical program if all approved parametric models do not fit the data and there is a reliable dose-response curve.

EXAMPLE of WET DATA ACCEPTANCE CHECKLIST

#	Acceptance Category	Data Validation Criteria	Action for Non-Conformance
1	Sampling	Was sample a 24-hour composite?	If "No," Retest with composite sample (or equivalent).
2	Sampling	Was sample taken from official permit compliance location (discharge outfall)?	If "No," Retest with sample taken at approved discharge location.
3	Sampling	Was sample bottle pre-rinsed 3-times prior to filling with effluent sample?	If "No," Re-test optional.
4	Sampling	Was sample bottle filled and sealed with minimal head space?	If "No," Re-test optional.
5	Sampling	Was sample temperature <4°C when it arrived at the laboratory?	If "No," Invalid sample; Re-test required.
6	Protocol	Was sample first used within the maximum allowed holding time (<36 hours after it was collected)?	If "No," Invalid sample; Re-test required.*
7	Protocol	Was dilution water chemistry within EPA specifications (alkalinity, hardness, conductivity, pH)?	If "No," Invalid test conditions; Re-test required.*
8	Protocol	Did organisms selected for inclusion in the toxicity test meet EPA age requirements?	If "No," Invalid test conditions; Re-test required.
9	Protocol	Did organisms selected for inclusion in the toxicity test meet EPA requirements for parental productivity?	If "No," Invalid test conditions; Re-test required.
10	Protocol	Were test organisms randomly-distributed according to EPA's recommendations for "blocking-by-parentage?"	If "No," Invalid test conditions; Re-test required.
11	Protocol	Did all test conditions comply with EPA's required protocols for temperature, dissolved oxygen, and feeding.	If "No," Invalid test conditions; Re-test required.*
12	Sensitivity	Was the most recent reference toxicant test within laboratory control limits?	If "No," Invalid test; Re-test required.
13	Sensitivity	Was the most recent valid reference toxicant test completed less than 30-days prior to the completion of the WET test?	If "No," Invalid test; Re-test required.
14	Sensitivity	Did the toxicity test meet EPA's minimum significant difference (MSD) criteria?	If "No," Invalid test; Retest required if MSD exceeded and test passed.
15	Sensitivity	Was the coefficient-of-variation for inter-replicate response among controls <40%?**	If "No," Invalid test; Re-test required.

** Deviations from some test conditions may be conditionally-accepted if approved by the permitting authority.*

*** North Carolina procedures as described in EPA's new WET guidance (June , 2000, appendix E & F)*

WHOLE EFFLUENT TOXICITY DATA ACCEPTANCE CHECKLIST *(continued)*

#	Acceptance Category	Data Validation Criteria	Action for Non-Conformance
16	Termination	Was the test terminated no less than 7 days or more than 8 days after the test was initiated?**	If "No," Invalid test; Re-test required.
17	Termination	Did 80% of control organisms produce at least three broods?**	If "No," Invalid test; Re-test required.
18	Results	Was reproduction calculated using only offspring from the first three broods?	If "No," Recalculate statistics using only first 3 broods.
19	Results	Was there a statistically-significant increase in mortality at all concentrations greater than or equal to the maximum permitted instream waste concentration (@99% confidence)?	If "No," then Not-Toxic If "Yes," probable toxicity when corroborated by a valid dose-response relationship.
20	Results	Was there a statistically-significant reduction in reproduction at all concentrations greater than or equal to the maximum permitted instream waste concentration (@99% confidence).	If "No," then Not-Toxic If "Yes," probable toxicity when corroborated by a valid dose-response relationship.
21	Corroboration	Is there a valid concentration response relationship confirmed by a statistically significant negative slope coefficient in a linear regression equation (@ 99% confidence)?	If "No," Sample is not certifiably toxic.
22	Corroboration	Was a statistically-significant increase in mortality corroborated by a statistically significant reduction in reproduction?	If "No," Inconsistent results; re-testing optional.
23	Corroboration	Do the NOAEC and IC-25 both confirm the presence of toxicity?**	If "No," Inconsistent Results; Report all results, unable to certify noncompliance.
24	Corroboration	If available, do identical split samples agree on the presence of toxicity at the maximum permitted instream waste concentration?	If "No," Inconsistent results; Report both tests, unable to certify noncompliance.
25	Anomalies	Was the mean control reproduction at least 15 but not more than 30 offspring per female?	If "No," Test out of control; Optional re-test.
26	Anomalies	Was the inter-replicate coefficient-of-variation abnormally low for control organisms?	If "Yes," Test out of control; Optional re-test.
27	Anomalies	Was the estimated IC-25 miscalculated due to bias introduced by required "data smoothing?"	If "Yes," recalculate IC-25 using 3-parameter logistic regression.
28	Anomalies	Was the reported toxicity test result likely to be an outlier as defined by the ASTM *h* and *k* statistics?	If "Yes," Optional re-test.