December 9, 2002
DRAFT

Summary of What EPA Has Done in its
WET Rule Published November 19, 2002

On November 8, 2002, the EPA Administrator signed a final rule — *Guidelines Establishing Test Procedures for the Analysis of Pollutants; Whole Effluent Toxicity Test Methods* — ratifying ten whole effluent toxicity ("WET") test methods and withdrawing two.[1] The Wet Coalition[2] submitted detailed comments on EPA's proposed rule. Our analysis of the final rule reveals that EPA has not corrected the serious flaws in the methods that we identified in our comments.

In the final rule, EPA ratified the (1) *Ceriodaphnia dubia* Acute Test; (2) Flathead Minnow Acute Test; (3) Sheepshead Minnow Acute Test; (4) Inland Silverside Acute Test; (5) *Ceriodaphnia dubia* Survival and Reproduction Test; (6) Fathead Minnow Larval Survival and Growth Test; (7) *Selenastrum capricornutum* Growth Test; (8) Sheepshead Minnow Larval Survival Test; (9) Inland Silverside Larval Survival and Growth Test, and (10) *Mysidopsis bahia* Survival, Growth and Fecundity Test. EPA revised some of the ratified methods "to improve performance and increase confidence in the reliability of the results." EPA concluded that these methods are (1) repeatable and reproducible, (2) available and applicable, and (3) representative, in accordance with the Agency's 1998 Report to Congress on the Availability, Adequacy and Comparability of Testing Procedures.[3]

---

[1] See 67 Fed. Reg. 69,952 (Nov. 19, 2002). The Agency originally finalized these methods on October 16, 1995, but failed to adequately address the extensive comments filed by stakeholders. In response, several industries filed a petition for review of the rule on February 21, 1996. On July 24, 1998, Petitioners and EPA entered into a settlement agreement.

[2] The WET Coalition consists of the following members: Alliance of Automobile Manufacturers, American Chemical Council, American Forest & Paper Association, American Petroleum Institute, AMSA, Rubber Manufacturers Association, Utility Water Act Group, VAMWA, WESTCAS, Alcoa, General Electric, Kennecott Utah, and Milliken Company.

[3] In our comments, we argued that EPA had not complied with the Report to Congress.

EPA withdrew the *Holmesimysis costata* Acute Test[4] and the *Champia parvula* Reproductive Test methods because it was unable to obtain interlaboratory precision data for them. In addition to these two methods, the WET Coalition recommended to EPA that the *Mysidopsis bahia* Fecundity Test must be withdrawn because EPA has not demonstrated by interlaboratory study that the proposed version of the method will eliminate the unacceptable performance problems identified during the interlaboratory study.[5] EPA did not withdraw the method because the fecundity endpoint is optional and not required for an acceptable test. EPA used this same rationale to justify not performing an interlaboratory study.

Below is a brief summary of the changes from the proposed rule and the Agency's response to comments it elected not to accept. A brief summary of the comments filed by the Coalition on these topics also is provided.

Section A. Changes from the Proposed Rule

1. Blocking by Known Parentage. EPA proposed requiring blocking by known parentage in the *Ceriodaphnia dubia* Survival and Reproduction Test method. EPA proposed to exclude from any reproduction analyses any "block" (replicates from the same parentage) where more than 50% of the surviving C*eriodaphnia* are identified as males. If less than 50% from a given block are male, only the males would be excluded.

The Coalition questioned this proposal because it could bias the test results due to very different approaches being used when the difference in percentages of males could be small (*e.g.*, 51% versus 49%) and appeared arbitrary. We recommended the exclusion of any block from reproduction analyses that contain males. EPA chose to modify the proposed changes to clarify that neonates from a single known parent may be used in the initiation of more than one test, but it retained the provisions we questioned.

2. pH Drift. EPA proposed specific procedures that could be used to control pH drift in chronic WET tests.

---

[4] While EPA withdrew the *Holmesimysis costata* test from Part 136, the test is "approved" for use on the Pacific Coast.

[5] EPA modified the method because of the low completion rate (only 50%).

The WET Coalition believed that EPA had not adequately addressed pH issues with its proposed revisions. The Coalition offered the following specific recommendations:

- include provisions for pH control in the acute method,
- require that a regulatory authority allow for pH controls during tests,
- provide language supporting pH control when the permittee has shown by parallel testing that pH drift or shift influences test results within the context of their use,
- communicate clearly in the methods the ability of pH drift or shift to contribute to toxicity rather than be the sole reason for toxicity,
- when tests are designed to address toxicity instream, set as the target pH the pH at the edge of the appropriate mixing zone; for chronic tests, the pH instream after dilution has taken place must be used rather than a pH found in the effluent stream.
- for end-of-pipe toxicity, maintain the pH at the pH of the sample upon completion of collection.

The Coalition commented on the confusion created by EPA's including in the methods the requirement that pH drift in the uncontrolled test be "substantially greater than" in the controlled test. The Coalition recommended that EPA remove the word "substantially" from the methods as it relates to pH control and that EPA also delete the sentences requiring pH drift in uncontrolled tests to be at least twice that of controlled tests.

The Coalition also objected to the use of the language providing regulatory agencies with the authority to request more information or additional testing before pH control is allowed and requested that EPA include in the methods the specific type and quantity of information required to justify pH controls in tests. Otherwise, regulatory agencies might use this test to justify unreasonable requests on permittees and labs. Finally, the Coalition requested that EPA support its statement that the daily cycle of upward pH drift and renewal may cause artificial toxicity (even in the absence of pH-sensitive toxicants) but only "in rare circumstances." If EPA cannot provide data to support the qualifier "in rare circumstances," it should remove it.

In the final rule, EPA did adopt some of the Coalition's recommendations and revised the procedures as follows:

- EPA clarified that, when the test objective is to determine the toxicity of an effluent in the receiving water, the target pH to maintain in a pH control test is the

pH of the receiving water measured at the edge of any mixing zone authorized in the permit;

- EPA clarified that, when the objective is to determine absolute toxicity, the target pH to maintain in a pH control test is the pH of the sample upon completion of collection;
- EPA clarified that in a pH control test the pH should be maintained within a ± 0.2 pH units of the target pH in freshwater chronic tests and within ± 0.3 pH units for marine estuarine chronic tests; and
- EPA added guidance on interpreting the results of parallel testing.

EPA also removed language from the proposed changes that warned about effects from pH drift in the absence of pH-dependent toxicants. The Coalition had recommended that EPA modify the language to make it clear that this is a concern, and not just "in rare circumstances," as indicated by EPA.

EPA declined to extend the proposed pH control guidance to acute test methods and failed to acknowledge the pH data referenced in the Coalition's comments. The Agency also pointed out that permittees could go to the added expense of continuous flow tests to address pH concerns.

3. Nominal Error Rates. EPA proposed to incorporate a nominal error rate of 0.01 instead of 0.05 for the sublethal endpoints for *Ceriodaphnia dubia* and fathead minnows. The Agency proposed the revision as required by the 1998 Settlement Agreement.

The Coalition commented that the lower error rate of 0.01 must be applied to all tests requiring the use of an error rate and pointed out that EPA had not provided a rationale for limiting its use. The Coalition also commented that EPA, by specifying that the lower error rate can be used only if the conditions in its "nominal error rate" guidance are satisfied, was proposing to convert that guidance into a binding rule without providing its technical justification for public comment.

In the final rule, EPA did not incorporate the proposed change in nominal error rate. Because no one provided any data to support changing the numbers for all tests and EPA had, in its opinion, no scientific justification for reducing the nominal error rate, EPA elected not to

reduce the error rate in any circumstances. In EPA's opinion, reducing the error rate reduces the power of the test (this is probably correct) and the chance of identifying toxic discharges, thereby reducing environmental protection.

4. Dilution Series. EPA finalized the proposed guidance on the selection of dilution series and WET testing and clarified that no one particular dilution series is required.

The WET Coalition commented that EPA should stop recommending a geometric dilution series because it adds uncertainty at higher concentrations. The Coalition supported the inclusion of the new language on selection of dilution for WET testing, but only with the addition of text specifying that the dilution selected must support calculation of defensible confidence intervals. This requirement must be mandatory for point estimate test results to be reliable.

EPA disagreed with the Coalitions comments. The Agency still recommends a geometric dilution series and did not add the requested text.

5. Dilution Waters. EPA finalized the proposed guidance on the selection of dilution waters and made minor modifications in response to comments to clarify that no single dilution water type is required for all tests. EPA clarified in the method manuals that "an acceptable dilution water is one which is appropriate for the objectives of the test, supports adequate performance of the test organisms with respect to survival, growth, reproduction or other responses that may be measured in the test, is consistent in quality and does not contain contaminants that can produce toxicity."

The Coalition supported the intent of EPA's new language on selecting a dilution water but recommended that the language be more flexible (i.e., the text of the guidance should make recommendations only on approach). While EPA did offer some flexibility with its modification, it still goes beyond simply making recommendations on approach.

6. Pathogen Interference. EPA finalized the proposed guidance on controlling pathogen interference in the Fathead Minnow Larval Survival and Growth Test. EPA clarified that it recommends pathogen control techniques that do not modify the sample. The Agency also provided "further explanation regarding the purpose for and required extent of pathogen source determination."

The Coalition commented that the method must provide a procedure to account for impact due to pathogens and be validated in multiple laboratories across the country. The Coalition also commented that EPA must provide language (1) supporting the use of defensible alternatives to those being proposed by EPA and (2) that the proposed procedures are not required prior to testing pretreated samples.

In response to comments about limiting pathogen interference, EPA clarified that the objective was to minimize interference "to the extent test results are not confounded by mortality due to pathogens." EPA dismissed the validation concerns because the technique is optional and the availability of data (generated by the Wisconsin Department of Natural Resources) to support it. EPA responded that the guidance contains a sufficient number of options and that permittees are not precluded from working cooperatively with the regulatory authority to develop other options.

7. EDTA and the *Selenastrum Capricornutum* Growth Test. EPA proposed to recommend the use of EDTA in the *Selenastrum* growth test, but not to require it. In the final rule, EPA now requires the addition of EDTA to nutrient stock solutions when conducting the growth test and submitting data under NPDES permits.

The Coalition commented that the test, both with and without EDTA, clearly is unreliable and must be removed entirely from the options for promulgated methods. The CVs (coefficient of variation) of 34.3% and 58.5% are unacceptable, even though in EPA's opinion, a CV of 34.3% is similar to chemical-specific methods and thus acceptable.

While a CV of 34.3% is not good, it is well within the range of other methods. The main concern with the EDTA method is the completion rate (61.4%).

8. Variability Criteria. EPA incorporated mandatory variability criteria for five chronic test methods (Fathead Minnow Larval Survival and Growth Test; *Ceriodaphnia dubia* Survival and Reproduction Test; *Selenastrum capricornutum* Growth Test; *Mysidopsis bahia* Survival, Growth and Fecundity Test; and Inland Silverside Larval Survival and Growth Test). EPA recommends the use of point estimation techniques over hypothesis testing approaches but did not make it a requirement. EPA is not requiring variability criteria for survival endpoints of acute methods because these methods are less variable than sub-lethal chronic test methods. EPA elected to "implement variability criteria (upper and lower "percent minimum sufficient

difference" bounds) as a test review step that is required when NPDES permits require sublethal WET testing endpoints expressed using hypothesis testing" instead of as a component of endpoint calculation.

The Coalition has serious concerns about EPA's use of percent minimum sufficient difference (PMSD) as described in our comments, the main one being that the 10^{th} percentile PMSD is no substitute for a WET detection limit. While EPA addressed some of our concerns in the final rulemaking, it did not address the overreaching issue of declaring an effluent toxic when it is not.

The final rule does not say that there is violation just because the variability if high (PMSD > the 90^{th} percentile), but it does require the test to be repeated if the 90^{th} percentile is not met. Tests that find significant differences even when the PMSD is in this high range are accepted. This may not be unreasonable considering that the power of the test is reduced when variability is high. For example, pathogens can often be a source of high variability that can lead to erratic tests which cause the effluent to appear toxic when it is not. EPA also did not provide the lower bound PMSD for all methods, which would be helpful, stating that it is not needed for tests with low variability.

EPA also failed to develop tools to address uncertainty in point estimates. But this is no longer necessary as EPA has dropped any consideration of data variability for point estimates. If EPA was to reconsider and take variability into account as the Coalition believes it should, setting limits on acceptable variability would be reasonable.

Given EPA's unwillingness or inability to develop detection limits for WET methods, one solution is to adopt the tiered approach in implementation. Under this approach, permitting agencies recognize that the methods are imperfect and should not be used to establish numeric permit limits. Instead, WET test results would be used to trigger additional testing and investigations to determine the cause of the apparent toxicity and to rectify it. A non-compliance would result from the failure to follow the procedure outlined in the permit, not from failing a WET test. Several states currently follow this or a similar approach.

9. Minimum Number of Replicates. EPA increased the minimum number of replicates in chronic fish and sea urchin tests and the *Ceriodaphnia dubia* chronic test as proposed. The Coalition supported an increase in the minimum number of replicates for a total

of four per control and concentration tested for the chronic fish and sea urchin tests. The Coalition did not support an increase of replicates in test methods (*e.g., Ceriodaphnia dubia* chronic test) that require four or more replicates without calculation of a defensible interlaboratory detection limit below which uncertainty of the data is too great to conclude a difference from control response.

10. Test Requirement Recommendations. EPA modified the tables of test conditions and test acceptability criteria to more accurately identify which test conditions are required, versus merely recommended. It also established two new requirements for WET test review: mandatory review of concentration-response relationship and mandatory variability criteria for some methods.

The Coalition expressed concern over the lack of clarity regarding what elements of a test method are mandatory rather than discretionary. The Coalition commented that EPA must revise the methods to indicate clearly what elements are mandatory. EPA appears to have done that. There is still a concern, however, as the enforcement agency, not the permittee, is the one who decides if a test is valid. Thus, a permittee could be in violation for failing to report an exceedance based on a test that was not performed in accordance with the method, but that the regulator deems acceptable.

11. Sample Collection and Holding Times. EPA clarified the requirements for sample collections and sample holding times by allowing samples to be used for renewal at 24, 48, and/or 72 hours after first use. The manuals provide flexibility when shipping renewal samples is delayed during an ongoing test. EPA also clarified that sample collections on days one, three, and five is the recommended, but not required, sample collection scheme. A minimum of three samples is required for seven-day chronic tests, but variations in the sampling scheme also are allowed.

12. Reference Toxicant Testing. EPA clarified the purpose and requirements for reference toxicant testing: (1) initially demonstrate acceptable laboratory performance; (2) assess the sensitivity and health of test organisms; and (3) document ongoing laboratory performance. EPA also clarified that "reference toxicant test results should not be used as a *de facto* criterion for rejection of individual effluent or receiving water tests."

The WET Coalition identified three flaws in EPA's reference toxicant testing procedure: (1) collection of long-term reference toxicant data is not mandatory; (2) the procedure does not promote reduction of variability within each laboratory; and (3) reference toxicant results that fall outside the control chart limit do not necessarily invalidate the associated test results. The Coalition also commented that, as reference toxicant test performance is a surrogate for WET test reliability, EPA must have an objective basis for what constitutes acceptable reference toxicant results.

EPA did add guidance for interlaboratory comparisons, but elected not to require the use of specific reference toxicants or to set required acceptance ranges. EPA also did not mandate the use of control charts, and results that are "out of control" are not sufficient to invalidate a test.

13. Sample Holding Temperature. EPA clarified the allowable sample holding temperatures for WET samples as 0-6° C. The WET Coalition had requested 2-6° C.

14. Biomass. EPA clarified that the sublethal endpoint used in survival and growth tests is based on the number of initial organisms exposed. The endpoint is a combined survival and growth endpoint, more accurately termed biomass.

The Coalition commented that EPA did not follow procedural requirements in adopting the biomass approach. Changing to a biomass endpoint does not make sense, but it does appear to make the test more sensitive.

15. Total Residual Chlorine. EPA clarified the requirements for measuring total residual chlorine in WET test samples. If total residual chlorine is not detected in effluent or dilution water at test initiation, it is unnecessary to measure it at test session renewal or at test termination. EPA retained the requirement to analyze total residual chlorine immediately following sample collection "because information on chlorine at the site and time of collection is important for evaluating the effectiveness of chlorination/ dechlorination processes and comparing the results of WET testing with instream effects."

The Coalition commented that the total residual chlorine testing should be either deleted or changed to apply only when the samples being prepared for testing or when the effluent being

tested has not been dechlorinated at the discharge site. EPA did not adopt either option, but the revised testing requirements seem less onerous.

16. *Ceriodaphnia dubia* Survival and Reproduction Test Determination Criteria. EPA modified the test to specify that neonates from the fourth brood on are excluded from the number of neonates counted in the reproduction test. Otherwise, EPA retained the current test termination criteria.

The WET Coalition recommended that EPA adopt North Carolina's additional test acceptability criteria for the *Ceriodaphnia dubia* reproduction test, one of which is the exclusion of neonates from four or more broods. This is the only criterion EPA elected to adopt.

17. Additional Minor Corrections. EPA made other additional minor corrections in response to comments.

Section B. EPA's Response to Comments

1. Validation of Performance Characteristics. The WET Coalition filed extensive comments on EPA's failure to validate the essential performance characteristics (accuracy, precision, dynamic range, detection limits, interference, ruggedness (applicability), reporting, and representativeness/method comparability) as laid out in the 1998 Section 518 Report to Congress. The Coalition pointed out that the methods are incapable of meeting minimum acceptable standards for these performance characteristics and that this is borne out by the evidence.

EPA disagreed with the Coalition's assertion that the WET test methods were not validated according to all the criteria identified in the 1988 Report to Congress. According to EPA, detection limits and dynamic range are not appropriate for WET test methods. EPA did apply the availability, adequacy, and comparability criteria identified in the report. In EPA's opinion, the methods verified were "available" because EPA has identified a sufficient number of laboratories that can conduct the test and culture the test organisms and "adequate" because a multi-laboratory test demonstrated a high degree of precision and the tests were reproducible.

EPA held to its position that the accuracy of WET methods cannot be measured. The Agency looked at accuracy as including both bias and precision, but stated that bias is not applicable to WET test methods because it cannot be described as it is for chemical analytes.

EPA believes it evaluated and considered the applicable performance characteristics (precision, interferences, ruggedness (applicability), reporting, and representativeness) in ratifying the WET test methods. EPA's response does not alter the Coalition's position that the essential performance characteristics were not validated. The Coalition pointed out that accuracy could be determined for blanks. Testing by Tim Moore indicated that changes in water chemistry might mimic toxicity in blank samples. EPA's interlaboratory study failed to test that hypothesis in its blank study — ampules instead of whole water samples were sent to participating laboratories. Nevertheless, EPA continues to tout the low false positive rate in blanks. More to the point is the fact that even if "true" toxicity is unknown, it is still possible to determine if there are systematic differences in the "amount" of toxicity measured by different laboratories. EPA made no attempt to measure those differences.

2. Interlaboratory Variability Study. The WET Coalition expressed concern that EPA considered test results to be valid when the laboratories did not comply with all the mandatory requirements of the test methods, including QA/QC requirements. The Coalition also pointed out that for two of the test methods (*Selenastrum* chronic and Silverside acute), EPA did not have the requisite number of laboratories (six) that were successful in completing the tests for all endpoints.

In response to comments that poor quality data were used from the interlaboratory feasibility study that would have been discarded in a regulatory context, EPA stated that (1) it appropriately evaluated the data according to the promulgated method and ASTM guidance for measuring interlaboratory precision and (2) it accurately invalidated tests according to the test acceptability criteria specified in each method. EPA agreed that it did not use the results of reference toxicant tests from the WET interlaboratory feasibility study to qualify or disqualify data. Instead, it used reference toxicant tests in the manner in which they are described in the method manuals.

EPA clarified in the final methods reference toxicant testing requirements and the appropriate use of the test data. EPA indicated that it did not establish precision criteria based on results from a single sample matrix. Precision estimates were based on the combined results of referenced toxicant, effluent, and receiving water testing on four sample matrices. Because multiple matrices were used to generate precision estimates, more than 6 usable data sets were

used for each method. Thus, according to EPA, more than six laboratories successfully tested all the methods it ratified.

3. Variability. The WET Coalition filed extensive comments related to the variability of the WET test methods. The Coalition rejected the argument that a coefficient of variation ("CV") of 0.26 was acceptable when measuring toxicity. We pointed out the lack of precision when EPA's acceptable range is plus or minus 100%. In response to EPA's claims that the WET test precision level is acceptable because it is comparable to chemical test method precision, the Coalition noted that the Agency's claim is true only if the best WET tests are compared to chemical methods operating at concentrations where imprecision is highest (i.e., at the detection limit). Moreover, the imprecision of chemical test methods is more tolerable because the results can be compared to a traceable standard to independently corroborate the accuracy. That option is not available for WET methods, where accuracy is unknown or unknowable.

EPA reiterated its opinion that the WET test methods are no more variable than other methods that are published in 40 C.F.R. Part 136 and used for regulatory compliance purposes. EPA rejected comments that the variability of the methods was too high for use in NPDES permits.

WET variability is at the heart of the contention between the Coalition and the Agency. EPA's guidance essentially says that permittees must accept the variability whatever it is, although the permittee may take steps to reduce it. More troublesome is EPA's position that statistical evaluation of the estimate by Point is not part of the compliance reporting. Thus, a permittee is considered to be in violation of its NPDES permit when the estimated concentration that causes toxicity, which is highly variable, is less than some limit even if that concentration is within the range of variation expected by chance.

4. Test Interference due to Method Requirements. The WET Coalition commented on various procedures that may unintentionally bias the results (e.g., test acceptance criteria to address test sensitivity and the ICp smoothing procedure).

EPA elected to retain the ICp method to calculate point estimates for sublethal endpoints. EPA did address the concern over ICp smoothing but did not acknowledge that there will be bias

any time control organisms are treated differently than the exposed organisms. Also, abandoning confidence intervals for compliance monitoring is unjustified.

5. Quality Assurance/Quality Control Requirements. The WET Coalition commented that the QA/QC criteria incorporated by EPA into the methods was inadequate and inconsistent with the 1988 Report to Congress and previous decisions the Agency had made under Part 136. Specifically, the Coalition commented that EPA's WET methods do not ensure comparability or representativeness of test results within and between laboratories. Nor do the methods adequately assess bias or sensitivity in test results within and between laboratories. The Coalition also commented that (1) EPA must adopt performance criteria that support regulatory use of WET data, (2) the QA/QC protocols must be mandatory, and (3) EPA must clearly state QA/QC requirements that define test validity.

EPA did not concur with the Coalition's assessment of the QA/QC provisions. EPA did add two additional QA/QC requirements, however: (1) review of concentration-response relationships and (2) mandatory variability criteria when NPDES permits require sublethal WET testing endpoints expressed using hypothesis testing. (These provisions are discussed in earlier sections.) The Agency believes that the current requirements are adequate to ensure reliable and quality results. As mentioned previously, EPA still does not require laboratories to keep control charts.

EPA indicates that, because variability criteria are assessed on an individual test basis, it is not necessary to set performance criteria for these measurements over time. Tracking over time is recommended but not required. EPA also recommends that laboratories compare their CVs to the CVs in the Agency's variability document.

The concern with EPA's approach is that it is not mandatory. Also, EPA does not address comparing the estimated values of an ICp endpoint between laboratories, only comparisons to variability. Different laboratories may estimate toxic concentrations for a specific reference toxicant that are 2 or 3 times higher or lower than some other laboratory. According to EPA, that variability is acceptable as long as it is within PMSD bounds.

6. Concentration-Response Relationships. The WET Coalition agreed with EPA's position that reviewing the concentration response is critical, but commented that it is not enough just to review the response. The absence or presence of a valid concentration-response

relationship must be confirmed before concluding that a sample is toxic. The Coalition also commented that it did not support EPA's guidance on concentration-response relationships. The guidance will not meet its intended goal and is biased to interpret the data in an unreasonably conservative fashion rather than in a scientifically supportable manner.

EPA requires the review of concentration-response relationships for all multi-concentration tests but does not require that a concentration-response relationship be established before determining that toxicity is present. EPA does not think establishing a valid concentration-response relationship before determining toxicity is appropriate for several reasons:

- WET methods and the WET testing program rely on the measurement of specific test endpoints, not on achieving specified concentration-response patterns.
- Concentration-response relationships are empirical. A single definition for a valid concentration-response relationship is not appropriate.
- WET testing has inherent characteristics that limit the ability to achieve ideal concentration-response relationships.
- Concentration-response relationship guidance has been shown to be very effective at reducing false positives.

The Agency also contends that the guidance is scientifically supportable and not unreasonably conservative. It included guidance for interpreting the concentration response when hypothesis is used.

EPA also failed to provide data on the slope of the concentration-response curves so that the impact of slope on the method sensitivity could be assessed. This type of analysis is needed to verify EPA's claim that the variability of WET methods is comparable to chemical methods. If that is true, it should be possible to estimate the concentration of a chemical by using WET methods.

7. Confidence Intervals. The WET Coalition commented on the need for confidence intervals to assess the reliability of results and to determine compliance.

EPA finalized its proposed method modifications that provide guidance clarifying what to do when the EPA software does not generate confidence intervals. But the Agency elected not to alter the compliance determination approach to include these confidence intervals. In EPA's opinion, it demonstrated in the interlaboratory variability study that these methods provide adequate precision and adequate protection from false positives.

In response to comments on the lack of confidence intervals in all circumstances, EPA states that the lack of confidence intervals should not adversely affect WET test result reporting because compliance is based on the point estimate itself. This is a clear reversal of past Agency policy when the basis for recommending the point estimate statistical procedure was the availability of confidence intervals.

EPA also takes the position that confidence intervals are irrelevant, even though the methods describe the ICp process for estimating confidence intervals. EPA has yet to publish guidance on the confidence interval problem. If such guidance is forthcoming, there may be an opportunity to influence it, depending on timing.

8. Dual Control Reporting. The WET Coalition commented on the lack of clear and correct reporting requirements for dual controls.

In response to comments on dual controls, EPA clarified the use of dual controls. Dilution water control is used to determine the acceptability of the test and for comparison with the tested effluent.

There does not appear to be any concept of an increase or net increase in toxicity in EPA's use of dual controls. Net increase in toxicity is an important issue, particularly for plants discharging non-contact cooling water.

9. Correlation with Instream Effects. The WET Coalition commented that EPA failed to validate the correlation between WET test results and actual instream conditions for all WET test endpoints. The Coalition also pointed out that the field validation studies do not demonstrate comparability of WET methods.

EPA disagreed with the Coalition's comments and pointed to several studies that demonstrate the correlation between WET test results (freshwater/marine and lethal/sublethal) and instream conditions. There is language in the preamble, however, that suggests EPA may

address means to adjust for frequency, duration, or magnitude of instream exposure conditions as an implementation issue.

Given the very high cost of field validating chronic endpoints and the remote chance that there would be any relationship between the chronic endpoint and instream effects, EPA is unlikely to undertake any studies to determine such relationships. To perform this type of evaluation, the study should consist of laboratory toxicity tests, ambient toxicity tests and community effects.

10. Method Flexibility. The WET Coalition expressed concern that "must" and "shall" provisions in the test methods, including QA/QC provisions, were not necessarily mandatory.

In response to comments that cut both ways, EPA attempted to balance the need for requirements versus the need for flexibility. (See previous discussion on Test Requirement Recommendations.)

11. Test Acceptability Criteria. As previously indicated, the Coalition recommended that EPA incorporate North Carolina's TAC for the *Ceriodaphnia dubia* test. In response to EPA's specific request for comments on increasing the test acceptability criteria for mean control reproduction in the *Ceriodaphnia dubia* Survival and Reproduction Test and mean control weight in the Fathead Minnow Larval Survival and Growth Test, the Coalition pointed out that EPA did not provide any justification for changing or a proposal for actual changes. Without such information, the Coalition could not provide meaningful comment.

EPA decided not to modify these tests because it received no data. EPA also elected not to establish mandatory variability criteria when NPDES permits require sublethal WET testing endpoints expressed using hypothesis testing. Instead, EPA incorporated these variability criteria as required test review steps for five methods. (See the previous discussion on variability criteria.)

12. Statistical Methods. The Coalition has serious concerns about the applicability of the statistical methods used by EPA and urged EPA to state explicitly that alternative approaches are allowed.

EPA elected not to include alternatives to statistical methods in the proposed test methods. It believes the statistical methods currently recommended are appropriate, but acknowledges that these test methods are not the only appropriate tests.

Guidance on using alternative statistical approaches would be useful. To get such guidance, the Coalition could recommend improvements to EPA's current approaches.

13. Successful Test Completion. The Coalition commented that EPA inappropriately claimed that a large percentage of the tests initiated as part of the large-scale interlaboratory validation study were completed successfully. Many of the tests were not performed in accordance with the method. In particular, the *Ceriodaphnia dubia* test completion rate indicates that it is not sufficiently robust to be ratified in Part 136.

In response to comments that successful test completion rates were incorrectly calculated in the WET interlaboratory variability study (failure to invalidate tests that did not meet specific test condition ranges), EPA stated that it accurately invalidated tests according to the test acceptability criteria specific to each method and successful test completion rates were based on meeting these criteria. EPA disagreed with comments that the *Ceriodaphnia dubia* test is not robust enough for inclusion in Part 136.

14. False Positive Rate. The WET Coalition commented that EPA's false positive rate of 5% (if accurate) is too high for the methods to be used for regulatory purposes. The Coalition recommended that EPA develop a detection limit for WET methods.

EPA disagreed with comments that false positive rates for WET test methods are unacceptably high. According to EPA, the interlaboratory variability study conclusively showed that measured false positive rates were below the theoretical rate of 5% estimated for the methods. EPA further indicated that the MDL concept was not applicable to WET test methods. In EPA's opinion, "the additional measurements, controls, replication and statistical approaches included in the WET test methods measurement system ensured that measured responses can be reliability distinguished from background noise."

EPA glories in the low false positive rate, which is not supported by Tim Moore's analysis of the data. EPA's censoring of the data, which was not an option available to permittees at the time, may explain the different results.

15. Age of Fathead Minnows in Acute Test. The WET Coalition commented that EPA's reduction in age range from 1-90 days to 1-14 days was not properly justified and not supported by recent data (Markle et al. (2000)).

EPA disagreed with the Coalition's interpretation of the Markle data and retained the age range of 1-14 days.

16. Outlier Analysis. The WET Coalition commented on the lack of guidance on how to manage outliers in general and for the *Champia* test in particular.

EPA responded that there is guidance (limited) on outliers in Appendix A of the 1994 chronic methods manuals. Regulatory authorities may provide additional guidance.

17. Cost. EPA estimated the total cost of the modifications for all permittees to be less than $5 million per year nationwide for all tests. EPA believes these costs would be alleviated by a potential reduction in costs for retesting and additional investigations. Thus, EPA did not see these costs as significant.

18. Implementation. EPA reiterated that the promulgation of the methods is not their implementation and that comments on implementation were not addressed, as they were not part of this rulemaking. The Coalition has met with EPA to discuss implementation issues and prepared a series of white papers that EPA is considering.

EPA perceived the Coalition's comment that the methods are not clear as to what endpoints are required as an implementation issue. The Agency did concur that the methods are not water quality criteria under the Clean Water Act.

In response to the Coalition's concerns regarding the stated primary objective of NPDES permit-related toxicity testing in section 7.1.1 of the acute and chronic methods, EPA removed the statement but points out that this is an implementation issue and should be handled in individual permit proceedings.