DEVELOPING A METHOD DETECTION LIMIT FOR CHRONIC WHOLE EFFLUENT TOXICITY TESTING USING *CERIODAPHNIA DUBIA*

Timothy F. Moore
Risk Sciences
1417 Plymouth Dr., Brentwood, TN 37027

Steven P. Canton
Chadwick Ecological Consultants, Inc.
5575 S. Sycamore St., Suite 101, Littleton, CO 80120

and

Max Grimes
Regulatory Management Inc.
6190 Lehman Dr., Suite 106, Colorado Springs, CO 80918

Abstract - Chemical testing routinely uses 'blanks' to provide quality assurance. However, whole effluent toxicity (WET) testing relies, primarily, on the use of reference toxicants. As such, the intrinsic variability surrounding WET testing in the absence of toxicants is not well known. For this study, a number of municipal wastewater dischargers contracted 17 laboratories to conduct a total of 25 chronic WET tests using the standard test organism, *Ceriodaphnia dubia*. Unbeknownst to the labs, the samples they received from the wastewater dischargers were comprised only of "moderately hard water" made using U.S. EPA's standard formula. As such, these tests served as 'method blanks'. Of the 25 tests completed by the biomonitoring laboratories, 2 did not meet control performance criteria. Of the remaining 23 valid tests, 9 (39%) indicated toxicity in the test sample (i.e., NOEC or IC_{25} < 100% 'effluent'). This failure rate was unexpected, considering the water being testing was identical among labs and comprised simply of mod-hard water. Using techniques similar to those employed for traditional chemistry, reproducible 'method detection limits' (MDLs) were calculated for the chronic *Ceriodaphnia* test. This calculation indicates that based on a standard 0.5 dilution series starting with 100% 'effluent,' at least 3 TU_c would be necessary to ensure that any reported toxicity was greater than the variability associated with this method.

KEYWORDS: *Ceriodaphnia dubia*, whole effluent toxicity tests, MDL, NOEC, IC_{25}

INTRODUCTION

Much of the debate surrounding whole effluent toxicity (WET) testing concerns the question of test variability. The U.S. Environmental Protection Agency (U.S. EPA) has stated that variability for standard WET test methods is within the range normally accepted for chemical analyses (see Federal Register, 40 CFR Part 136, October 16, 1995). Chemical testing compensates for such measurement error by defining a method detection level (MDL). To date, no MDL has been developed for whole effluent toxicity test methods. Without this important tool, WET test variability may have a much greater influence on NPDES compliance determinations than does the analytical variability also found in chemical analyses.

Federal regulations define a method detection level as that concentration at which the technician is 99% confident that the analyte is truly present when the test says it is present and 99% confident that the analyte is truly absent when the test says it is absent (U.S. EPA Method 40 CFR 136.2). In other words, it is the point at which the test can accurately and reliably distinguish the difference between signal and noise.

U.S. EPA has maintained that "accuracy of toxicity tests cannot be ascertained only the precision of toxicity can be estimated" (see Federal Register, October 16, 1995), acknowledging that biological organisms are inherently variable. However, unlike chemical analyses, there is no way to calibrate a living creature to a known standard when conducting a WET test.

Previous attempts to quantify the variability of WET testing focused primarily on the response of test organisms to reference toxicants [1- 5]. These results from identical reference toxicant samples, analyzed simultaneously at different laboratories, have been compared and used to estimate WET test precision. A similar effort has not been conducted in the absence of toxicants.

While it may not be possible to accurately assess the magnitude of toxicity present, it is possible to evaluate whether the test can reliably distinguish between signal and noise. If a WET test failed to identify a known toxic sample as "toxic," then that would be a false negative. Likewise, if a test failed to identify a known non-toxic sample as "non-toxic," then that would be a false positive result. Both are errors in accuracy (Table 1). Because all reference toxicants are toxic by definition, they cannot be used to estimate the occurrence of false positives. The true incidence of Type-I errors can only be appraised using non-toxic samples. In chemical testing, this is called a "method blank."

Table 1: Accuracy in Whole Effluent Toxicity Testing

	REFERENCE TOXICANT	METHOD BLANK
TEST FAILS	Accurate: True Positive	Inaccurate: False Positive
TEST PASSES	Inaccurate: False Negative	Accurate: True Negative

Studies conducted with reference toxicants suggest that WET tests perform well at detecting the presence of toxic pollutants [6]; i.e., there are relatively few false negatives. This does not imply, however, that the WET test procedures are equally accurate at affirming the *absence* of toxic pollutants. Evaluating the presence or absence of toxicity is not like viewing two sides of the same coin. For example, a smoke alarm may accurately forewarn the presence of fire in a building. If, however, the same smoke alarm frequently announces the presence of fires which do not exist, it is inaccurate.

To date, no WET test studies have been conducted using method blanks. However, recent retrospective analyses using results from control organisms suggest that there is considerable biological variability even in the absence of known toxicants [7, 8]. Statistical analyses indicate the level of biological variability may cause the actual occurrence of false positives to be 2-3 times higher than expected [9, 10]. Monte Carlo simulation studies have provided some empirical data to support these conclusions [8].

Since WET test results are being used to certify both the presence and the absence of toxic pollutants in wastewater discharges, then it is essential to know the true rate of false positives and false negatives. To do so, reference toxicant studies must be supplemented by data from method blank studies. This paper describes the first prospective method blank research on whole effluent toxicity testing.

METHODS

The objective was to prepare a large volume of non-toxic water and analyze identical aliquots at several different bioassay labs simultaneously. The aliquots were made to look like ordinary effluent samples and the labs were unaware that they were part of a quality assurance evaluation using the *Ceriodaphnia dubia* chronic test methods.

A total of 375 L of standard synthetic freshwater was prepared. The source water had been filtered through activated carbon, de-ionized three times, and final-filtered through a 5 um media, resulting in 18 megaohm water (equivalent to ASTM Type I water). Salts and minerals were added in conformance with U.S. EPA's recommended formula for moderately hard water [11]. Chemical analysis showed that it was within the specified ranges of pH, conductivity, alkalinity, and hardness (Table 2). Subsequent chemical analyses performed by the WET laboratories also confirmed that the non-toxic sample water was formulated correctly (Table 2).

Table 2: Chemical Characterization of Synthetic Freshwater Samples

Parameter	Range Specified in EPA's Guidance for Ceriodaphnia Tests	Measured Concentration *(internal QA/QC)*	Average Reported Value *(measured by Participating labs)*
Alkalinity	60-70	66 mg/L	68 mg/L
Conductivity	n/a	308 umhos/cm	309 umhos/cm
Hardness	80-100	92 mg/L	91 mg/L
pH	7.4-7.8 s.u.	7.7 s.u.	7.78 s.u.

Additional chemical analyses were performed to assure that the water was not contaminated by toxins (Table 3). It was particularly important to know whether the water quality was changing while stored in the sample containers. For this reason, several Cubitainers® were selected at random and held in cold (4° C) storage for later chemical analysis over time periods meant to mimic the time period for shipping to the dischargers and their subsequent shipping to the bioassay laboratories.

Samples from the containers were analyzed after the initial preparation (Sample 1), on the first day the dischargers sent samples to their WET laboratories (Sample 2), and on the last day samples were shipped to the WET laboratories (Sample 3) (Table 3). Based on the analyses of these samples, there was no evidence that there were any toxins present in concentrations sufficient to cause toxicity in a WET test using the standard invertebrate species (Table 3). This same source water is routinely used by a certified chemistry lab to calibrate analytical instrumentation. In addition, the source water has a well-documented history as the basis for preparation of appropriate dilution and control water in WET testing in our laboratory, including successful completion of all U.S. EPA's DMR-QA studies since 1991.

The synthetic freshwater was prepared in five new 110 L Nalgene® containers that were thoroughly rinsed and conditioned prior to use. Samples were dispensed into new Cubitainers® from a single outlet using new tubing and gang-valves, which were also well-rinsed and prepared under laboratory-controlled conditions.

Seventy-five Cubitainers® were filled, packaged, and shipped via overnight courier to participating wastewater agencies. The samples were immediately refrigerated upon receipt. Each discharger re-labeled the containers as "effluent," "reference toxicant," "process control samples," or simply with a numbered sample code. Each agency had arranged, in advance, to conduct a chronic WET test at the bioassay laboratory they traditionally used for permit compliance monitoring.

Table 3: Water Quality Analyses for Synthetic Freshwater Samples

PARAMETER	SAMPLE 1 (Day 1)	SAMPLE 2 (Day 5)	SAMPLE 3 (Day 10)
Chloride	1.6 mg/L	1.5 mg/L	1.7 mg/L
TDS	170 "	190 "	180 "
Sulfate	74 "	78 "	71 "
Calcium	14.0 "	14.5 "	14.1 "
Magnesium	13.3 "	13.8 "	13.8 "
Potassium	1.9 "	2.0 "	2.1 "
Sodium	29.3 "	30. "4	31.5 "
Cadmium	0.6 ug/L	<0. 6 ug/L	<0.6 ug/L
Chromium	<5 "	<5 "	<5 "
Copper	<5 "	<5 "	<5 "
Iron	<30 "	<30 "	60 "
Lead	<3 "	<3 "	<3 "
Manganese	<1 "	<1 "	<1 "
Mercury	<0.2 "	<0.2 "	<0.2 "
Nickel	<5 "	<5 "	<5 "
Selenium	<5 "	<5 "	<5 "
Silver	<0.5 "	<0.5 "	<.5 "
Zinc	<5 "	<5 "	<5 "
EPA 524 Organics (59 compounds @ 0.5-1.0 ug/L)	ND	Not tested	ND
EPA 608 Organics (18 compounds @ .003-.083 ug/L))	ND	Not tested	ND
EPA 8260 Organics (56 compounds @ 5-10 ug/L))	ND	Not tested	ND
EPA 8270 Organics (144 compounds @ 10-50 ug/L))	ND	Not tested	ND

The labs were instructed to conduct a standard whole effluent toxicity test using U.S. EPA's most recent chronic method for *Ceriodaphnia dubia* [40 CFR 136 Method 1002.0, 11]. They were told to use U.S. EPA's mod-hard recipe water for preparation of dilutions and controls during the test. The labs were asked to perform a concurrent reference toxicant test and to certify full conformance with U.S. EPA's recommended procedures. The labs were not told that the samples were part of a quality assurance study.

Samples were sent from the dischargers to the biomonitoring labs in three separate shipments (Monday-Wednesday-Friday) beginning September 29, 1997. This is consistent with the every-other-day manner in which effluent is normally collected and transported during WET testing. It also helped obscure the true nature of the samples. All WET testing was finished by October 7, 1997, well within the two week holding time recommended for synthetic freshwater [11].

Laboratories were asked to run the tests using a standard 0.5 dilution series with six treatments from 100% to 0% "effluent." The labs were also instructed to provide copies of all raw data from the bench sheets. They were to calculate test results using the most appropriate statistical procedures as specified in U.S. EPA's data analysis flowchart [11], specifically providing both the NOEC (hypothesis-testing) and IC_{25} (linear interpolation) values. These were then used to calculate toxic units (TUc = 100/NOEC or $100/IC_{25}$). Non-toxic samples would be expected to have a TUc = 1. Therefore, any result reported as greater than 1 TUc was considered a false positive.

RESULTS

A total of 25 tests were initiated at 17 different labs (Table 4). Some labs were contracted by more than one wastewater facility, however, no lab performed more than 3 tests. All 25 tests met U.S. EPA's minimum acceptance criteria for control survival (>80%). And, no test showed a statistically-significant difference in mortality when sample-exposed organisms were compared to controls. Thus, there were no false positives when 7-day *Ceriodaphnia dubia* survival was used as the test endpoint.

Two of the 25 tests (8%) failed to meet U.S. EPA's minimum acceptance criteria for reproduction (Table 4), which requires control organisms to average at least 15 offspring per surviving female. These two tests were discarded from further analysis. The remaining 23 chronic tests (92%) met U.S. EPA control performance criteria and were deemed valid for subsequent analysis.

When evaluated by test, 14 of the 23 valid tests (61%) showed no statistically-significant reduction in reproduction when sample-exposed organisms were compared to controls (Table 4), whether the endpoint was calculated based on NOEC or IC_{25}. These results are consistent with the initial assumption that U.S. EPA's formula for moderately-hard dilution water is not toxic. However, 9 of the 23 valid tests (39%) indicated the presence of toxicity by reporting an NOEC or an IC_{25} less than 100 (Table 4). Estimated toxicity was reported in a range from 2 TUc to more than 16 TUc.

When evaluated by laboratory, 2 of the 17 labs (7%) failed to meet U.S. EPA's control performance criteria (and the results rejected, as noted earlier). Nine of the 17 labs (53%), conducting 14 tests, observed no toxicity based on reproduction, while 6 of the 17 labs (35%), conducting 9 tests, reported that the method blank samples were chronically toxic to *Ceriodaphnia dubia* based on the sublethal endpoint of reproduction. All labs that conducted multiple tests reported consistent results among their tests.

Table 4: Chronic Ceriodaphnia Reproduction Results Using Method Blanks

Test ID	Control Survival	Sample Survival	Survival Pass/Fail	Control Reprod.	Sample Reprod.	Reprod. NOEC	Reprod. IC25	Reprod. Pass/Fail
A	100	100	Pass	30.0	33.4	100	100	Pass
B	80	80	Pass	26.0	26.4	100	100	Pass
C	100	90	Pass	28.5	28.0	100	100	Pass
D	100	100	Pass	16.4	17.0	100	100	Pass
E	90	100	Pass	18.0	19.1	100	100	Pass
F	100	80	Pass	15.8	14.4	100	100	Pass
G	100	100	Pass	37.4	42.7	100	100	Pass
H	100	100	Pass	36.1	41.2	100	100	Pass
I	90	90	Pass	21.5	29.1	100	100	Pass
J*	90	NA	Pass	30.9	N/A	>17	>17	Pass
K	90	100	Pass	27.9	31.1	100	100	Pass
L	90	100	Pass	21.6	22.3	100	100	Pass
M	100	100	Pass	19.7	16.2	100	100	Pass
N	90	100	Pass	15.6	17.2	100	100	Pass
O	100	90	Pass	18.0	8.2	25	31.9	**Fail**
P	100	90	Pass	17.5	6.8	50	69.8	**Fail**
Q	100	80	Pass	18.1	7.8	50	46.5	**Fail**
R	100	80	Pass	21.5	7.4	50	37.9	**Fail**
S	100	90	Pass	34.4	27.5	50	100	**Fail**
T	100	100	Pass	29.8	13.1	25	50.8	**Fail**
U	100	80	Pass	32.1	18.9	6.25	36.7	**Fail**
V	80	80	Pass	15.5	9.1	100	11.6	**Fail**
W	90	80	Pass	15.1	10.9	100	12.1	**Fail**
X	90	90	Pass	12.2	9.0	*6.25*	*3.3*	*QC Reject*
Y	100	100	Pass	11.5	7.6	*100*	*43.9*	*QC Reject*

**Note: test "J" was not run with a .5 dilution series. The highest tested concentration was 17% of the original synthetic freshwater sample. In addition, receiving water was used for controls and dilution rather than the moderately-hard formula specified by written instruction.*

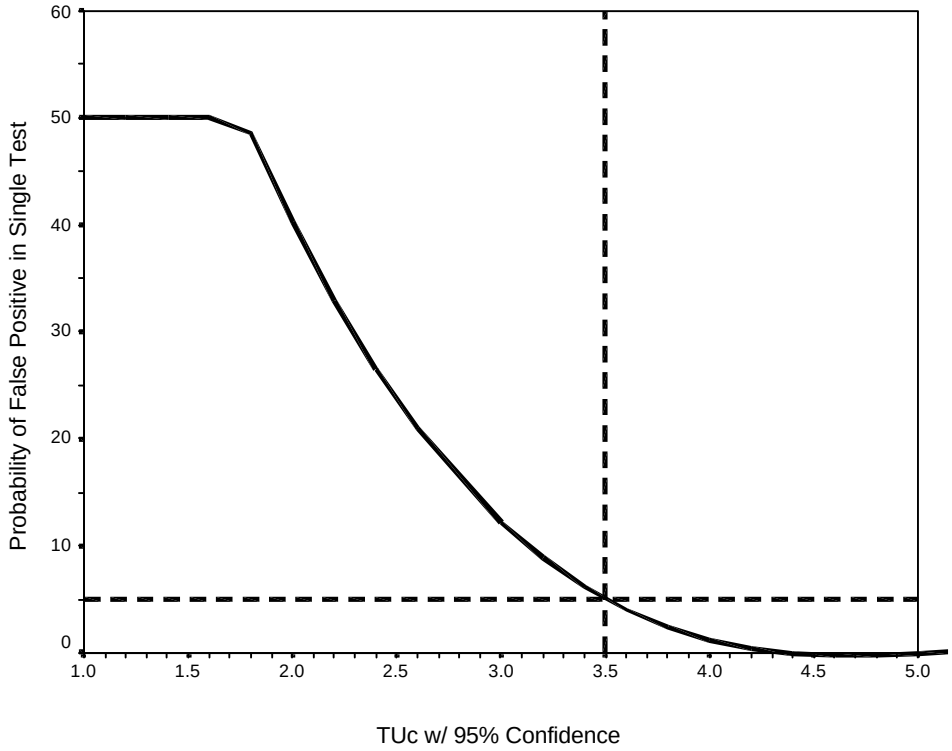
DETERMINATION OF METHOD-DETECTION LEVEL

As with most biological studies, WET tests rely on statistics, which include an inherent risk of error. U.S. EPA guidance recommends the use of 95% confidence intervals when analyzing toxicity data. Therefore, the established error-tolerance for WET testing is one false positive in every 20 test (5%).

Because all of the labs performed their statistical analyses using alpha = 0.05, only one of the 23 valid tests would have been expected to record a false positive. The actual incidence of Type-I errors was seven times higher than expected. This suggests that assigning a low alpha-value does not, by itself, restrict the occurrence of false positives to the acceptable range. As such, calculation of a method detection level for *Ceriodaphnia* reproduction would be appropriate, as outlined below.

The first approach for developing a method detection level would be to use the same techniques relied on in chemical analyses, as described in ASTM's new Standard Practice for 99/95% Interlaboratory Detection Estimate (IDE) for Analytical Methods with Negligible Calibration Error [D6091-97; as cited in Water Environment Laboratory Solutions, April/May 1998, p. 9]. For this approach, seven blank samples are analyzed to establish the normal range of background "noise" in the test. The mean and standard deviation are used to estimate the 95th percentile of the signal/noise ratio.

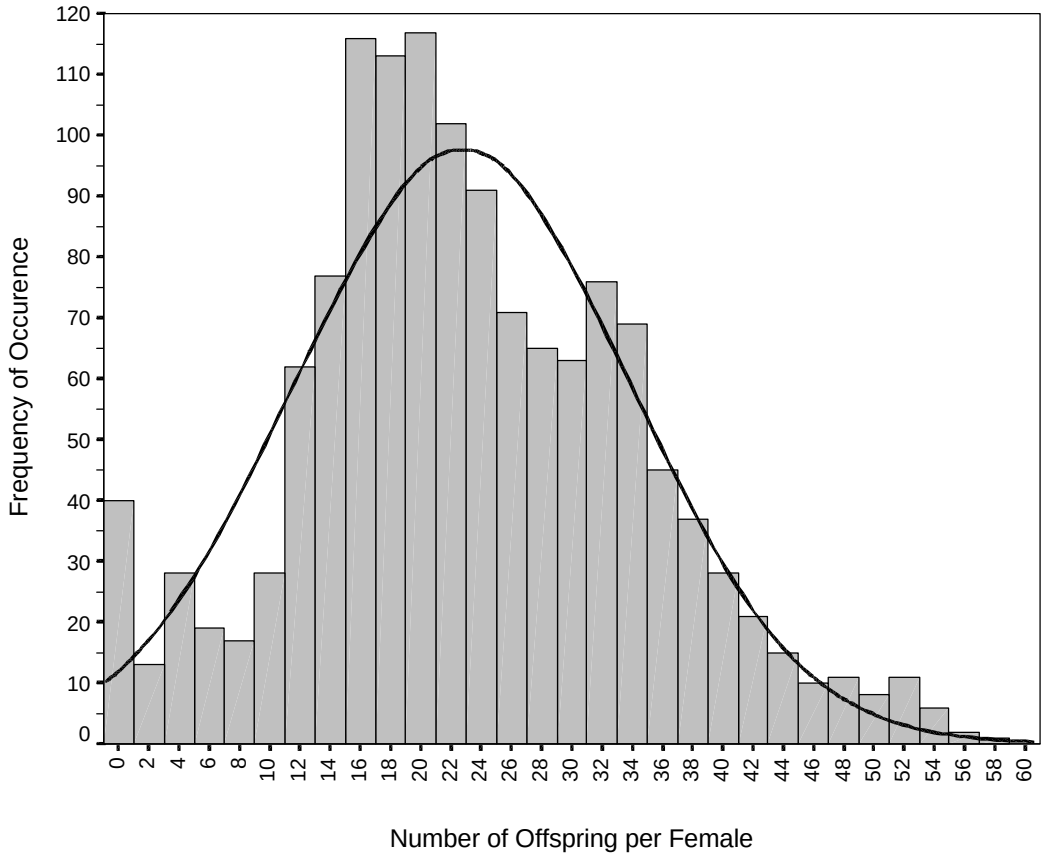
Figure 1: Estimated MDL for Ceriodaphnia dubia Reproduction



It would be very difficult and expensive for each bioassay lab to run seven additional blank test series each time they conducted a WET test. However, it is possible to estimate the approximate MDL using re-sampling techniques to analyze data from the method blank results in Table 4. Monte Carlo analyses were conducted simulating repeated sets of seven WET method blanks (based on the 0.5 dilution series ranging form 0-100%). This analysis indicated that the 95th percentile is approximately 3.5 TUc (Fig. 1), or 28% "effluent." That is, based on these data, we cannot be 95% confident that toxicity is actually present until the final test estimate (NOEC or IC_{25}) exceeds 3.5 TUc or less than roughly 30% "effluent" (Fig. 1, above).

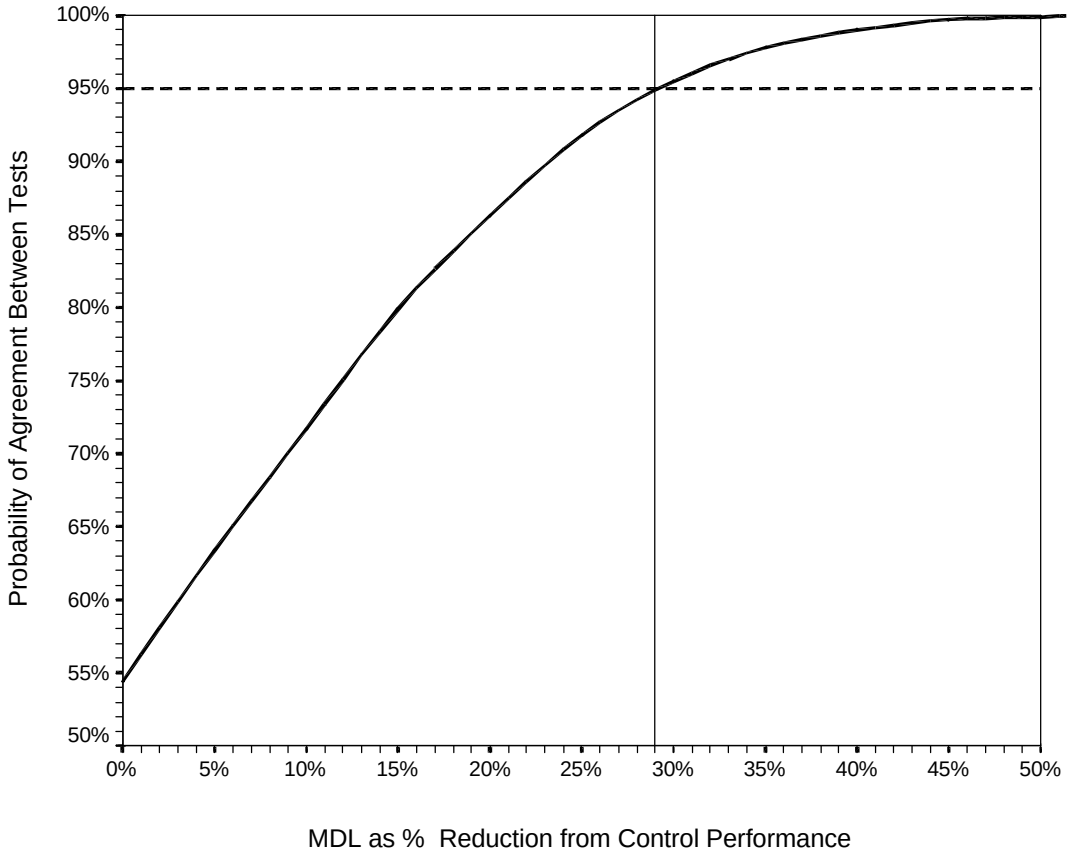
A second approach for calculation of an MDL for WET would be to estimate the smallest reduction in reproduction that WET tests can reliably distinguish from zero under non-toxic conditions. Data from the method blank study shows that the true mean reproduction for individual *Ceriodaphnia* is approximately 22.2 offspring per female (Fig. 2), with a true standard deviation of approximately 11.8 offspring per female (based on the 1,500 replicates in the study). In other words, given these data, the average reproduction for randomly selected groups of ten organisms will vary between 15 and 29 offspring per female 95% of the time.

Figure 2: Range of Ceriodaphnia dubia Reproduction Observed in Non-Toxic Water



Monte Carlo simulations using these data indicate that the average difference in mean reproduction between two randomly selected groups of ten organisms will be less than 30% of the control performance 95% of the time. The re-sampling analysis suggests that smaller reductions in reproduction occur too commonly among known non-toxic samples to conclude that toxicity is actually present based on a minimum magnitude of effect and a minimum confidence level (Fig. 3). Based on this approach, toxicity would be presumed to be absent, until we were 95% confident that there was at least a 30% reduction in the test endpoint of reproduction. Conceptually, this is similar to the IC_{25} with it's bootstrap confidence intervals.

Figure 3: Estimated MDL for Ceriodaphnia dubia Reproduction Using EPA's Linear Interpolation Method (IC-25)

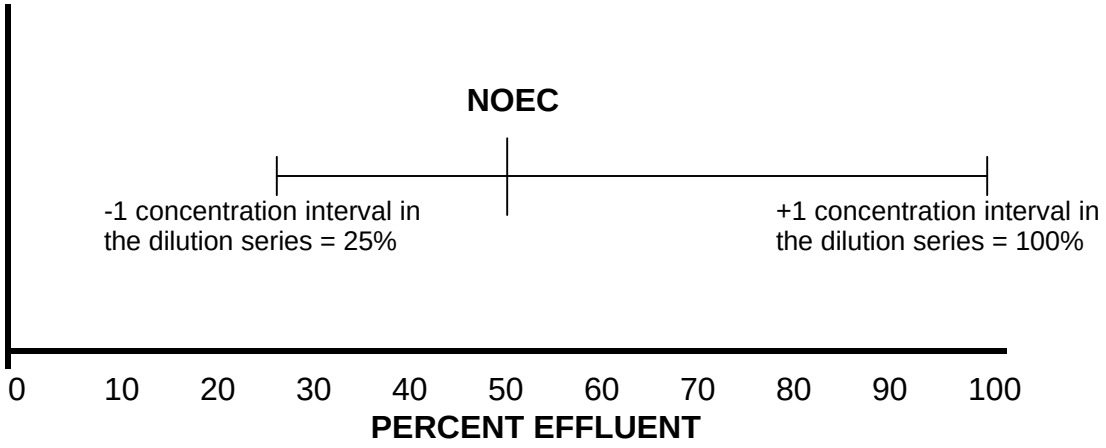


Note, figure 3 is read as follows: there is a 95% chance that two labs, each analyzing identical aliquots, will agree that toxicity is "present" (e.g. >1 TUc) if IC-25 values less than 30% are reported as "non-detect." The labs will not necessarily agree on the magnitude of toxicity present. Agreement regarding the magnitude of toxicity would require a PQL be established.

A third approach, independent of our method blank study, would be to calculate an MDL consistent with the level of inter-test repeatability identified by U.S. EPA for the IC_{25} point estimates. Federal guidance states that the expected level of precision for *Ceriodaphnia* reproduction is a 40-70% coefficient-of-variation [4, 11]. This implies a standard deviation which is approximately half the mean value. It also suggests that the 95th percentile should fall in a range that is approximately plus or minus 100% of the mean estimate of toxicity. Therefore, if the true answer for a non-toxic sample is 1 TUc, and one expects a coefficient-of-variation of 40%, then it is impossible to confidently conclude that toxicity is present until the reported estimate is at least 2 TUc. If the coefficient-of variation is closer to 70%, then the effective method detection level is approximately 3.3 TUc, similar to that shown in Figure 1.

A fourth approach, also independent of the method blank study, can be based on the fact that when the WET methods were finalized, U.S. EPA stated that the expected error band for inter-test precision using NOEC was "plus or minus one concentration interval in the dilution series" (see Federal Register, October 16, 1995, p. 53535). This means that a test result with a reported NOEC of 50% may more accurately be described as an estimated range between 25% and 100% (see Figure 4).

Figure 4: Recommended "Error Band" for WET Test Precision When NOEC=50%



This conclusion is confirmed by U.S. EPA's interlaboratory studies. Each year, U.S. EPA conducts a nationwide survey of bioassay labs using reference toxicants as part of their DMR-QA evaluation program. Results from these reference toxicant studies show a wide variation in reported values for identical samples (Table 5 & 6). The actual 95% confidence range from these data is approximately plus or minus 1.5 dilution intervals from the median toxicity estimate. If each concentration interval represents a 0.5 dilution factor, then the effective MDL suggested by the DMR-QA data is approximately 3 TUc - again, closely matching that calculated from the method blank study.

Table 5: Variation in Reported NOEC Value Using EPA Reference Toxicants

Ceriodaphnia dubia Reproduction as NOEC	N of Labs	Median Value	2.5th Percentile	97.5th Percentile
EPA DMR-QA #12 (1992)	103	20.0	1.60	50.0
EPA DMR-QA #13 (1993)	124	25.0	6.25	50.0
EPA DMR-QA #14 (1994)	147	25.0	6.25	50.0
EPA DMR-QA #15 (1995)	147	25.0	6.25	50.0
EPA DMR-QA #16 (1996)	140	25.0	6.25	50.0

Table 6: Variation in Reported IC-25 Value Using EPA Reference Toxicants

Ceriodaphnia dubia Reproduction as IC25	N of Labs	Mean Value	2.5th Percentile	97.5th Percentile
EPA DMR-QA #12 (1992)	101	24.1	4.9	47.6
EPA DMR-QA #13 (1993)	128	20.3	3.2	60.0
EPA DMR-QA #14 (1994)	141	27.5	6.3	48.8
EPA DMR-QA #15 (1995)	145	27.2	5.6	49.7
EPA DMR-QA #16 (1996)	141	28.2	8.1	51.0

DISCUSSION

Given the surprisingly high number of false positives during the method blank study, it was prudent to suspect that the synthetic freshwater might have been contaminated with toxins. In fact, regression analysis did confirm a potential concentration-response relationship when all of the valid data results are pooled ($p<0.008$, $R^2=0.005$, $N=1362$). However, despite the significant regression, the relationship accounted for less than one-half of one percent of all the observed variance among replicates. This would indicate that the test water was no sufficiently toxic to cause the number of false positives observed.

In addition, if the synthetic freshwater were truly slightly toxic, then one would expect the most sensitive tests to identify the phenomena first. That was not the case. There was a statistically-significant positive correlation between test sensitivity and the probability of *passing* the test ($p<0.033$, $R^2=0.59$, $n=23$). This also indicates that the test water was not toxic.

While the results from the method blank study were surprising, the performance by the test organisms (mean = 22.2 offspring, std. dev. = 11.8, n = 1,500) was consistent with data reported by other researchers. The first large-scale interlaboratory analysis of WET test precision found wide variability for *Ceriodaphnia* reproduction in non-toxic water, with reproduction ranging from 0 to 55 offspring per female among control organisms [12]. Similar variability (range = 0-50+ offspring, mean=21-22 offspring) was reported by Oris and Bailer [7], Chapman et al [13] and Canton et al [8].

SUGGESTED IMPROVEMENTS

This high MDL and the number of false positives reported herein could be reduced by several mechanisms. First, Type-I error can be controlled directly by reducing the critical alpha value from 0.05 to 0.01 when performing statistical analyses of WET data [14]. U.S. EPA now permits such adjustments for chronic toxicity test endpoints, such as reproduction for *Ceriodaphnia* or growth for the fathead minnow (*Pimephales promelas*) test [Settlement Agreement, 24 July 1998, in Edison Electric Institute et al vs. USEPA, U.S. District Court of Appeals, D.C. Circuit, Case No. 96-1062].

Second, this study indicates that increased reliance on mortality endpoints, with reduced reliance on sublethal metrics, would reduce the influence of biological variability on compliance determinations. Although reproduction and growth have been considered "more sensitive" indicators, the data indicates that this perceived increased sensitivity is not actually present. Recent studies suggest that mortality endpoints also provide more reliable indicators of instream impairment [15].

Third, functional MDLs can be established to more reliably distinguish true toxicity events from biological background noise. Such an approach would be consistent with U.S. EPA's recommendation that “the allowable frequency for criteria excursions should refer to true excursions of the (numeric or toxicity) criteria, not to spurious excursions caused by analytical variability or error” [4]. This study suggests that such an MDL would be approximately 3 to 3.5 TUc. Alternatively, MDLs could be calculated by the U.S. EPA by their insertion of method blanks in their annual DMR-QA reference toxicant samples.

Finally, U.S. EPA could require a labs to run a full-dilution series method blank at least once each month. Long-term method blank control charts, similar to those currently required for reference toxicant performance, would be maintained for the method blank results. Data from the method blanks could be used to calculate laboratory MDLs and potentially identify methods to improve laboratory procedures. In addition, dischargers could routinely supplement their own QA/QC programs by submitting blind method blanks to the biomonitoring lab.

Some states have already adopted some or all of the above recommendations [16] as part of a more comprehensive effort to establish and meet data quality objectives for whole effluent toxicity testing.

The implications of this research for NPDES permitting are significant. Regulatory decisions, such as reasonable potential analyses, compliance monitoring, toxicity identification evaluations, and 303(d) listings, are currently based on WET test results that do not take into account an MDL that may be 3 or greater. If WET test results are to be used like chemical analyses to assess permit compliance, then a WET MDL must be employed to preserve the credibility of the results.

REFERENCES

1 DeGraeve GM, Cooney JD, McIntyre DO, Pollock TL, Reichenbach NG, Dean JH, Marcus MD. 1991. Variability in the performance of the seven-day fathead minnow (*Pimephales promelas*) larval survival and growth test: an intra- and interlaboratory study. *Environ. Toxicol. Chem.* 10:1189-1203.

2 DeGraeve GM, Cooney JD, Marsh BH, Pollock TL, Reichenbach NG. 1992. Variability in the performance of the 7-d *Ceriodaphnia dubia* survival and reproduction test: An intra- and interlaboratory study. *Environ. Toxicol. Chem.* 11:851-866.

3 Anderson SL, Norberg-King TJ. 1991. Precision of short-term chronic toxicity tests in the real world. *Environ. Toxicol. Chem.* 10:143-145.

4 U.S. Environmental Protection Agency. 1991. *Technical Support Document for Water Quality-based Toxics Control* (TSD). EPA 505/2-90/001, Washington, DC, USA.

5 Dorn PB. 1995. An industrial perspective on whole effluent toxicity testing. pp. 16-37. IN: Grothe DR, Dickson KL, Reed-Judkins DK (eds.). *Whole Effluent Toxicity Testing: An Evaluation of Methods and Prediction of Receiving System Impacts*. SETAC Press, Pensacola, FL.

6 Grothe DR, Dickson KL, Reed-Judkins DK (eds.). 1995. *Whole Effluent Toxicity Testing: An Evaluation of Methods and Prediction of Receiving System Impacts*. SETAC Press, Pensacola, FL.

7 Oris JT, Bailer AJ. 1993. Statistical analysis of the *Ceriodaphnia* toxicity test: Sample size determination for reproductive effects. *Environ. Toxicol. Chem.* 12:85-90.

8 Canton SP, Dantin DD, Moore TF. 1998. *Chronic Whole Effluent Toxicity Test Organism Performance, with Notes on WET Test Statistics. Environ. Toxicol. Chem.* (In review).

9 Dhaliwall BS, Dolan RJ, Smith RW. 1995. A proposed method for improving whole effluent toxicity data interpretation in regulatory compliance. *Water Environment Research* 67:953-963.

10 Warren-Hicks WJ, Parkhurst BR. 1995. Issues in whole effluent toxicity test uncertainty. pp. 180-189. IN: Grothe DR, Dickson KL, Reed-Judkins DK (eds.). *Whole Effluent Toxicity Testing: An Evaluation of Methods and Prediction of Receiving System Impacts*. SETAC Press, Pensacola, FL.

11 Lewis PA, Klemm DJ, Lazorchak JM, Norberg-King TJ, Peltier WH, Heber MA. 1994. *Short-Term Methods for Estimating the Chronic Toxicity of Effluents and Receiving Waters to Freshwater Organisms*, Third Edition. EPA/600/4-91/002. U.S. Environmental Protection Agency, Environmental Monitoring and Support Laboratory. Cincinnati, OH.

12 DeGraeve GM, Cooney JD, Marsh BH, Pollock TL, Reichenbach NG. 1989. Precision of the EPA seven-day *Ceriodaphnia dubia* survival and reproduction test: Intra- and inter-laboratory study. EPRI EN-6469. Electric Power Research Institute, Palo Alto, CA.

13 Chapman GA, Anderson BS, Bailer AJ, Baird RB, Berger R, Burton DT, Denton DL, Goodfellow Jr WL, Heber MA, McDonald LL, Norberg-King TJ, Ruffier PJ. 1995. Methods and appropriate endpoints: Discussion synopsis. pp. 51-78. IN: Grothe DR, Dickson KL, Reed-Judkins DK (eds.). *Whole Effluent Toxicity Testing: An Evaluation of Methods and Prediction of Receiving System Impacts*. SETAC Press, Pensacola, FL.

14 Denton DL, Norberg-King TJ. 1995. Whole effluent toxicity statistics: A regulatory perspective. pp. 83-102. IN: Grothe DR, Dickson KL, Reed-Judkins DK (eds.). *Whole Effluent Toxicity Testing: An Evaluation of Methods and Prediction of Receiving System Impacts*. SETAC Press, Pensacola, FL.

15 Daley C, Barbour M, Diamond J. 1998. Defining relationships between whole effluent toxicity testing and instream toxicity. *Proceedings Surface Water Quality and Ecology Symposium*, Water Environment Federation, Arlington, VA.

16 Washington State Department of Ecology. 1997. Laboratory Guidance and Whole Effluent Toxicity Test Review Criteria. Publication No. WQ-R-95-80.