

Responsiveness Summary on Comments Received from the Review of the Draft 2005 Canary Book (Ecology Pub. # WQ-R-95-80)

Comments and Questions:

The roles and authority of the Water Quality unit and the Environmental Lab Accreditation Program (ELAP) are not completely clear, particularly for allowing modifications and exceptions to the test methods. If a method modification was part of the SOP approved by the ELAP, can it be assumed the modification would also be acceptable to the Water Quality unit? If both organizations are involved in approving method modifications, should the revised SOP be sent to the ELAP **and** to the Water Quality unit?

The canary book describes how whole effluent toxicity (WET) test reports are reviewed. Each test type has a summary checklist in the canary book that is used in these reviews. Detailed SOPs are not used in the test report reviews. SOPs are reviewed and approved solely by the ELAP.

Need some discussion, general comment and guidance on the alternative tests, such as the herring and trout embryo viability tests. - just a suggestion that background information may be provided in the form of an Appendix.

Discussions will be included in the final canary book of the herring tests, the trout embryo test, and the trout 7-day survival and growth test.

p. 11 Section II. B. 6. Sample Hardness Adjustment:

In adding reagents to the sample to adjust it to the hardness of receiving water or storm water, is there any guidance as to tolerances for the ultimate hardness relative to the receiving water/sample (as, $\pm 5\%$ or whatever)? And is there any provision for aging the water once hardness has been adjusted? (In preparation of reconstituted water, EPA methods call for allowing 24 hours to dissolve the added chemicals and stabilize the medium. Would this have any application in the Hardness Adjustment?)

We have reviewed the draft of the Dept. of Ecology Publication No. WQ-R-95-80, *Laboratory Guidance and Whole Effluent Toxicity Test Review Criteria* and we have a few comments regarding section B.6. Sample Hardness Adjustment. You recommend adjusting the hardness of the effluent to match the receiving water, it is our understanding that the effluent should not be manipulated. EPA guidance only refers to adjusting the hardness of the dilution water to match the receiving water if receiving water is not used in the test. Wouldn't adjusting the dilution water, rather than the effluent, be a more accurate representation of what is occurring in the system?

Adjusting the dilution water hardness to match the receiving water would cause the test concentration series to better simulate the effluent as it combines with receiving water downstream of the discharge. However, the purpose for allowing hardness adjustment is to avoid hardness gradients across test concentrations when the sample has low hardness. Adverse effects due to low hardness will increase in the same direction in the concentration series as sample toxicity and defeat using the concentration-response relationship to screen out effects not due to toxicity. Low hardness can cause the death of test organisms, reduce *Ceriodaphnia* reproduction, make fathead minnows more susceptible to infection, etc. Please keep in mind that all samples tested with marine organisms are salted up to 30 ppt and tested without a salinity gradient which might confound the results.

EPA recommends aging reconstituted water for 24 hours before using it in toxicity tests so that its constituents can fully dissolve and equilibrate. There will not be enough time in most circumstances to add the hardness constituents directly to a sample, wait 24 hours, and start the toxicity test within the 36-hour holding time. The canary book will be revised to include two methods to compensate for this problem. For all acute WET tests and any chronic test for a discharge without a defined point of compliance (ACEC and CCEC are unknown.), the test should include a test chamber with 100% sample that has not had hardness added and a second control made from deionized water that has had hardness added to match the receiving water at the same time that the hardness constituents were added to the sample. If a chronic test is to be performed on a sample from a discharge with a known ACEC and CCEC, then a high hardness water may be prepared a day or more in advance and used to bring the hardness of the sample up to the receiving water hardness as long as both the ACEC and CCEC can be included in the test concentration series and the test includes a control consisting of deionized water to which enough of the high hardness water has been added to equal the receiving water hardness.

The receiving water hardness must be equaled in the test concentrations to within plus or minus 10%.

Section B, part 1: Transfer and Storage:

The End Time is to be recorded in Pacific Standard Time. All other times recorded in the lab use either standard or daylight savings time, whichever is in effect. To follow a convention for recording the end time for composites that is inconsistent with all other time values (during daylight savings time) could be problematic. Since the sample time is critical only in how it compares to the time the sample as received and when the test was started and ended, these should all be recorded on the same basis.

The requirement to use Pacific Standard Time for recording sample times is simply a recognition that the State of Washington is in the Pacific Standard Time zone. This is not in conflict with “daylight savings time” which is always relative to the local time zone. The requirement was added because some dischargers record sample times relative to work shifts and argue with the overworked lab who must take the time to translate this unofficial time convention into PST. We are aware that inconsistencies can occur and cause miscalculation of sample holding time if the lab is in a different time zone than the discharger or if daylight savings time has gone on or off while the sample is in transit. Labs should include a mention in the test report of any inconsistency of this type.

Collecting and storing samples with no headspace is required. With glass containers, headspace is present during the compositing period and difficult to completely eliminate when the compositing is completed and during storage for renewals. It may be a more practical approach to require storage with *minimal* headspace rather than *no* headspace. Perhaps an allowable option would be storage under a nitrogen headspace, which is allowed for sediment bioassays.

The document will be revised to encourage the use of a nitrogen headspace with glass containers.

Section C, part 1: Randomization:

In addition to randomization of test chamber positions as per EPA methods, there is a requirement for randomization of bench sheets (for weight measurements). The current EPA methods describe approaches for organism distribution and test chamber positioning but I wasn't able to find an approach for randomization of measurements. It would be helpful to include in the manual a randomization technique for bench sheets and measurements if this is a method requirement.

The EPA manuals do not provide for randomization during the test beyond the randomization of test chamber positions. The randomization of test chamber positions should be reflected on the bench sheets. Properly done, all measurements, including weights, will be done in the random order reflected on the bench sheet by the test chamber position randomization.

Section F. Reference Toxicant Tests:

Routine reference toxicant test results should be reported monthly. Should the non-routine reference toxicant tests (done every 6 months as an accreditation requirement) also be sent as a separate report, even if they are included with the sample report? It would be helpful to have additional details, particularly in the Electronic Submission of Test Data as to how routine and non-routine reference toxicant data should be submitted if reported outside a sample report.

The reports on reference toxicant tests done once every six months as an accreditation requirement should be held for review during ELAP audits.

Fathead Minnow Survival and Growth Test (p. 36):

The DOE protocol says to feed the larvae 3x/day or 2x/day at set brine shrimp (Artemia) concentrations. We feed the larvae 4x per day (sometimes less the first day of the test to prevent buildup of uneaten food waste and DO drop/ammonia when the larvae are young) to meet survival and growth requirements. During method development for this test, we tried various feeding regimes, and have test data that documents better larvae performance with the 4x/day feeding. The DOE Lab Accreditation Program has accepted our method.

The same feeding regime issues apply to the Mysidopsis bahia, Topsmelt (Atherinops) and Silverside (Menidia) Survival and Growth Tests (pp. 41, 42 and 43, respectively). For the Topsmelt test, care is taken

not to overfeed the larvae. For the *Mysidopsis* test, cannibalism is an issue, so larvae are fed a slight excess of *Artemia*.

The document as written allows fathead minnows to be fed 4 times/day at 2.5 - 3.0 hour intervals. The commenter should make sure that they are reviewing the current draft. The mysid, topsmelt, and inland silverside feeding instructions merely reflect what is said in the EPA test manuals. The feeding instructions in the canary book should be viewed as guidance. Caution is urged but labs may use other feeding routines. No change in methodology should be made with revising the SOP first and seeking approval from the ELAP. The inability to produce topsmelt 7-day survival LC50s < 205 µg/L Cu as described in the next comment could be due to overfeeding.

Topsmelt (Atherinops) Survival and Growth Test (p. 42):

The DOE states that an acceptable LC50 endpoint value must be less than 205 µg/L for a copper chloride reference toxicant test. However, out of seven reference toxicant tests run so far, three were indeterminate (no observed effect; LC50 > 245 µg/L) and the most recent two resulted in an average LC50 of 210 µg/L. The EPA and DOE requirement is unreasonable. It is possible that larvae from a different supplier might give a more sensitive response, but we prefer a supplier that provides larvae with proven health and performance.

Why is an LC50 used as an acceptance criterion in the Topsmelt ref tox test? Isn't an IC25 ± 2SD more appropriate?

A brief survey of the database and labs has determined that some labs can keep 7-day LC50s below 205 µg/L copper and others cannot. The LC50 < 205 µg/L acceptance criterion is not a practical requirement at this time. The EPA manual expresses a preference for copper chloride but does not forbid the use of other reference toxicants and this could be used by labs as a way around the LC50 < 205 µg/L Cu acceptance criterion anyway. The LC50 < 205 µg/L Cu acceptance criterion will remain in the canary book as a goal and copper chloride will be specified as the only acceptable reference toxicant while information on the acceptance criterion is gathered for the next revision of the canary book. Labs all use the same test organism source so this is not the reason for differences. Differences in feeding, test organism loading, salinity, and/or pH are more likely to be the cause. EPA may have developed this acceptance criterion in tests using hypersaline brine and will be asked for more information. The usual QC plotting of reference toxicant test results as described in the EPA manual and elsewhere in the canary book still apply to topsmelt. The canary book will be revised to make clear that the reference toxicant test requirements on this page are in addition to the other standard requirements.

p. 1 Section I. B:

I would suggest adding something about labs getting hold of a copy of the permit in order to know exactly what it requires. Labs have commented to me that clients are sometimes hesitant to show them the permit. If the client is reticent to share a copy of the permit (even after the Lab shows them the Canary Book statement that the Lab should have access to the permit) the Lab may request a copy from the Dept of Ecology since it is public information (there may be a small fee).

The document will be revised to include this advice. It is possible that permittees could get angry if they did not want a lab to have a copy of the permit and the lab went directly to the Dept. of Ecology to get one anyway.

p. 6 Section II. B. 1., First and second paragraphs:

Under Federal Register 40 CFR Part 136, Vol 67, No. 223 (November 19, 2002) and the three WET testing method manuals of 2002 (EPA-821-R-02-012, EPA-821-R-02-013, and EPA-821-R-02-014, sections 8.6.1), acceptable sample holding temperatures was changed from 4°C to 0-6°C. I agree that 4°C is a great target temperature, but are we deviating from the EPA Manual requirements on this one? This same question applies to II. B. 2., as well.

The commenter is correct and the document will be changed to be consistent with EPA.

p. 16 Pathogens:

Do they ever originate from the source of purchased test organisms? I thought we once had a discussion where you mentioned that that was happening in some labs.

No, I did research using our database on whether the source of fathead minnows had any effect on sporadic mortalities assumed to be caused by pathogens or on final weights. Neither sporadic mortalities nor final weights show any relationship to fathead minnow source. In addition, the controls in most labs have little or no mortalities in tests with sporadic mortalities indicating that the sample might be the source of any pathogen. I have publicly speculated that test organism suppliers have raised a tremendous number of generations of fathead minnows under disease-free conditions and that this has contributed to their susceptibility to pathogens in samples.

p. 23 Section II. F. Fifth paragraph:

In the last sentence it states that "...this information can be useful when deciding the consequences of reference toxicant testing conducted by the lab." I'm not sure I know what that means.

The table below that paragraph describes when WET tests are acceptable or not based on reference toxicant test results. The results of reference toxicant tests by test organism suppliers can sometimes help get WET tests accepted even though testing lab reference toxicant results are outside of control limits. For this reason, labs are encouraged by that paragraph to use test organism suppliers who do routine reference toxicant testing even though test organism supplier reference toxicant results cannot substitute for testing lab reference toxicant testing. Consult the table for more details.

p. 24 Table, second column, second row:

"If the 95% confidence intervals can be calculated *and in both* failing ..." Is that what you intended to say? You lost me.

That row in the table describes the decisions made when a reference toxicant test and its first repeat are both outside of control limits. If the 95% confidence interval for the point estimates from these reference toxicant tests overlap the control limits in the QC plot, then a WET test might be accepted if other conditions described in the table are also met. The 95% confidence interval for a point estimate from a reference toxicant test result is the range which has a 95% probability of containing the "true" value of the point estimate. If the 95% confidence interval overlaps the control limits in the QC plot, then it can be considered that there is a reasonable chance that the "true" point estimate might have been within control limits. More explanation of this will be added to the document.

p. 27 Section II. G. 2. Temperature, last paragraph:

It seems to me that the first sentence may be a problem. It comes across as if you're minimizing the significance of two very important issues in bioassay testing. Labor savings and reduced potential for cross-contamination of test chambers are not trivial matters in a tox lab!!!

Also, it may be worthwhile to mention that "dummy" test replicates should be the same as the test vessels and volumes, in order to accurately reflect the temperatures of adjacent "real" test replicates. That way temperature changes will be at the same rate in both, reflecting changes in the surrounding temperature of the water bath or the incubator/environmental chamber.

Labor savings and minimizing cross-contamination are not trivial, but the paragraph does describe how the benefits from using a surrogate test chamber for monitoring temperature are not as great as a lab might think. For example dissolved oxygen and pH are measured with probes and must be done in the real test chambers regardless of whether surrogates are used for temperature. The sentence will be reworded to appear less negative about labor saving and cross-contamination. Another sentence will be added emphasizing the need for all surrogate chambers to closely resemble test chambers.

p. 28 Section II. H. third paragraph:

I suggest the first sentence read "...a lab must *contact* Randall..." to make it agree with the phone number and e-mail address.

OK

p. 52:

I would like to recommend that the reference to the Biomonitoring Science Advisory Board that was previously in Appendix B, be returned to this section in some form. It gave a bit of authority to the

discussion that I think is important. As it is now, it comes across as arbitrary decisions that were made more or less unilaterally.

Appendix B was shortened and simplified considerably to reflect changes in the canary book over the years that simplified the approach to bivalve endpoint calculation. Bivalve endpoint calculation evolved from just using proportion normal (BSAB days), to using complex combined endpoints (early EPA West Coast manual), to a complex process to determine the most sensitive endpoint among 3 possibilities (result of a multi-state and EPA meeting in Portland, OR to discuss bivalve endpoints), and finally to a decision to simply use both proportion alive and proportion normal. Proportion normal and proportion alive are simple calculations with a long history of use in science and environmental regulation. Some effluents affect survival more than development and some effluents do the opposite. Combined endpoints obscure this information. The decision was based on good information and is easier to implement than the more complex approaches in past canary books. Earlier versions of the canary book and Appendix B are still available for those interested in the history.

Monitoring for Changes in Toxicity, page 3:

- Comments on the steps to follow in case a rapid screening test has failed, permittees without a WET limit, and non-compliance. I suggest that for clarity, a step-by-step process (some flow-through chart) would be appropriate.

Permits will describe rapid screening test requirements and the response to toxicity. Because not all permits will be written the same, a detailed description in the canary book might give a misimpression and certainly would distract labs and permittees from consulting permits for the specific requirements.

Aeration, page 10:

- When dissolved oxygen level is a problem, there is a chance the DO could decrease well below the 4.0 mg/L for justification of aeration, and may have already killed the organisms prior to aeration. To avoid this, a pre-emptive aeration may be an option, if there is a good chance that the DO will fall below the 4.0 mg/L based on a trend (i.e., monitoring the DO several times during the day).

The current language was not intended to prevent pre-emptive aeration of test chambers. A sentence will be added to make this clear.

Appendix C, page 53:

- Additional comments on calculation of growth or combined survival and growth discussion. My opinion is that growth response is a secondary endpoint and it is appropriate to look at the growth response when there are no significant mortalities in the test treatments relative to the negative control. The growth response is less critical when there is already significant survival response. There is really no need to consider the applicability of growth endpoint when the organisms are already dying.
- Calculating growth based on the final survival may also result in anomalous response and also difficulties in coming up with some point estimates. For example, the mean calculated response in the test treatments sometimes may show the same mean weight or even higher than the control. Based on some past observations, there were times when there were very few surviving fish (most likely the strongest) and actively feeding in the sample treatments, therefore, resulting to normal-sized fish. If the growth of this fish is analyzed based on the final survival, the weights will most likely be the same as the control, consequently resulting in no detectable growth adverse effects or inability to calculate point estimates.

While it is true that survival is the most important measured endpoint in a toxicity test, EPA and the State of Washington are also interested in the sublethal endpoints when they are the more sensitive response in a test which also has reduced survival. Our default is the combined weight and survival (“biomass”) endpoint which equals the weight/initial count and we only switch to the original growth endpoint (weight/final count) when variability makes the biomass endpoint unreliable. The only situation requiring a point estimate here is calculating reference toxicant test results for use in QC plotting. Combined survival – sublethal endpoints can be used for QC plotting. A sentence will be added to make this clear.

Page 6, last paragraph. Glass containers are not necessarily superior to plastic – for samples in which metals are a concern, it would be preferable to use plastic in my opinion.

Permittees and labs may use plastic sample containers. The language will be edited to not favor glass so strongly.

Page 10. Sample Aeration. I have an alternative suggestion for the approach with respect to treating samples with supersaturation. This is perhaps not as frequently a concern in Washington as it is in BC, but in cases where samples are collected at colder temperatures than the test temperature, supersaturation is not unusual. Because the full-strength sample is the most supersaturated, the agitation associated with preparing dilutions does not really assist with reducing the problem of supersaturated dissolved oxygen - the fact that it takes some time to prepare dilutions does help to provide time for equilibration; however, since the saturation level is being approached asymptotically, it actually would be expected to take some time to reach this level without mixing of some sort. Most laboratories are able to prepare dilutions quite quickly, so this does not provide much time for equilibration. Preparing and then pouring dilutions, and then having to aerate every test chamber, also results in a substantial additional effort for the laboratory.

An alternative approach that I have found useful in cases where saturation is exceeded is to warm the sample to a few degrees above the test temperature, and then aerate or stir the sample on a stirplate. Raising the temperature above test temperature results in lowering the saturation asymptote below saturation at test temperature, and results in a more rapid decline in DO levels.

Supersaturation is not much of a problem here anymore. The alternative approach sounds good but is too involved to incorporate into the document without greater need. Labs may use it nevertheless.

Page 14. Suppression of pH Rise. 2nd paragraph, 1st line, I'd suggest wording it "try aerating with CO₂-supplemented air", rather than "try aerating with CO₂ mixed with O₂", and similarly on the 3rd line of that paragraph, "a slow flow of CO₂-supplemented air over test chambers". I wouldn't recommend aeration with an air that does not include nitrogen, because this could result in substantial oxidation and potentially oxygen supersaturation.

OK

Page 23, 3rd para, last line, after "the consequences of", I'd suggest adding "results associated with"

OK

Page 35. The first sentence appears to be a requirement (a must) – in fact, I would imagine that this would not necessarily be done by a laboratory, and the second sentence of this section provides an example of when this might not be done. Suggest changing the first line to a "may", or joining the first two sentences with a "however"

The sentences have been combined for clarity.

FYIs:

- Accreditation of tox labs is by the species name used in the current EPA Method Manual. And since only the most current manual is written into compliance monitoring permits, that is the only EPA reference for which accreditation is granted. Some labs have moved to use of the *Americamysis* designation. A few are using the *Raphidocelis* species name. But all know that the EPA designation will be used for accreditation until the new names are in the EPA manuals.
- The ballast water biocide testing protocol was revised to include internationally recommended tests, more detail on test solution renewals, and a description of circumstances when herring toxicity testing is needed.
- A list of preferred usages of CETIS fields was included. Labs that follow the list are voluntarily standardizing usage so that data can be shared between users without sacrificing full database capability.

Compliments and Endorsements:

Other than that section, everything looks good to us, and is very clear.

As with earlier versions of this manual, this version is a good guide to the use and interpretation of the WETT methods and WAC requirements.

Overall, the document is precise and clear, and provides an excellent source of technical guidance.

p. 19 Section II. E. 2. Changing the *alpha* for small differences in response. This section is so well written...it's a complex issue, but your wording makes it very understandable. Bravo!