External Peer Review of Draft "Aquatic Life Water Quality Criteria for Selenium – 2002"
Response to Comments from

Dennis Lemly
U.S. Forest Service, Southern Research Station
Coldwater Fisheries Research Unit
1650 Ramble Road, Blacksburg, VA 24060

April 11, 2002

Notice: EPA intends to reexamine the peer review issues after the general public has had an opportunity to provide information, data, and views on the 2004 draft document. Note also that unless otherwise noted, all references to page numbers or appendix letters are for the 2002 draft. The 2004 draft has more pages of main text and more appendices than the 2002 draft.

General Comments: The document is well researched, generally well written, and I agree with the approach of using a fish tissue-based method to derive the criterion for chronic exposure. I also agree with both the Final Acute Value (185 µg/L) and the Final Chronic Value (7.9 µg/g dw) for new national criteria.

Response to Questions in the Technical Charge:

Acute Criteria in Fresh and Salt Waters

Comment: 1. The toxicity tests used to derive the criterion are appropriate for that purpose. I am not aware of other relevant data that were not used for the acute criterion assessment.

Response: So noted.

Comment: 2. The acute criteria are appropriate for their intended purpose, that is, direct waterborne exposure for short durations.

Response: So noted.

Comment: 3. I believe that the strength of the relationship for sulfate influence on selenium toxicity is sufficient to support expressing the freshwater acute selenate criterion as a function of sulfate concentrations.

Response: The revised draft document includes a sulfate correction for the selenate FAV that was derived similar to the hardness correction for certain metals. Appendix A of the revised draft document presents the data and analysis that was used to determine the correction.

Chronic Criterion

4. Is a concentration in whole-body fish tissue an appropriate basis for expressing the criterion?

Comment: The selenium concentration in whole-body fish tissue is an appropriate basis for expressing a criterion for chronic exposure because it integrates the 3 major pathways for selenium uptake (water, planktonic food chain, benthic-detrital food chain), and encompasses the most sensitive biological endpoints (fish reproduction and teratogenic effects).

Response: So noted.

5. Are the toxicity tests and other studies used to derive the criterion appropriate for such use? Are you aware of other relevant data that were not used?

Comment: The toxicity tests and other studies used to derive the criterion are appropriate for that purpose. I am not aware of other relevant data that were not used for the chronic criterion assessment.

***Response*:** So noted.

6. Is the freshwater chronic criterion appropriate?

Comment: The freshwater chronic criterion is appropriate if it is understood that site-specific modifications will likely be necessary under some circumstances.

Response: We recognize that site-specificity may call for a greater or lesser level of protection than that provided by the FCV. Language was added to the revised draft document to allow for the development of a site-specific FCV where appropriate.

7. With the goal of being neither under- nor overprotective, how reliable would you expect the criterion to be in application to different sites? Are there any straightforward ways of improving its site specificity?

Comment: Given that the acceptable level for impacts is an EC_{20} (pg. 47), the criterion should provide adequate protection for most species of fish most of the time. However, if Centrarchids have whole-body tissue residues near the criterion level (7-8 µg/g dw) concurrently with winter conditions (cold temperature and short photoperiod), then unacceptable (>20%) mortality of juveniles may occur due to Winter Stress Syndrome. In this situation, which could be widespread, and probably others yet to be identified, site-specific revision of the criterion will be necessary. However, such modifications can't be done from a tissue-basis alone. Even though EPA has chosen to "sidestep the controversy involved in setting a reliable water concentration" (Versar background document, pg. 2), the point of regulation for selenium will still be water,

whether promulgated as a criterion by EPA, as it has been in the past, or stepped down to states and tribes to deal with on their own, as it will be now. Thus, as a purely practical matter, a national tissue-based criterion will not eliminate the need for states and tribes to set water concentration-based limits on selenium sources, and also develop and implement TMDLs to reduce selenium to acceptable levels in fish tissues. The tissue criterion is a biological target but the mechanism for meeting the target will still be manipulating waterborne selenium. States and tribes will have to decipher what waterborne concentration (standard) is required to keep whole-body residues in fish at or below 7.9 µg/g dw. This necessitates developing new (and/or re-examining old) state- or site-specific, water-based standards. Thus, formulating selenium standards is now a more circuitous process – EPA gives the biological target and it is up to states/tribes to find the waterborne concentration necessary to achieve it, which could be more difficult and time-consuming than the old (1987) water-based criterion technique. I don't disagree with the tissue-based approach, but it is important that EPA not leave states and tribes on their own to wrestle with the issue using trial and error methodologies to set a water standard. With the switch to a tissue-based criterion, proper guidance for site-specific modifications is more necessary than ever before. Because of selenium's unusual ecological risk factors (complex aquatic cycling pathways and multiple modes of toxicity), implementation guidance for this trace element must be selenium-specific in order for states and tribes to formulate appropriate standards. EPA's current implementation guidance is woefully inadequate because it is generic – it does not provide the necessary degree of specificity. However, in response to this information need (a need that I identified long before the EPA Peer Consultation Workshop or this Draft Criteria Document), I have developed and published a peer-reviewed procedure for deriving site-specific chronic criteria for selenium. The method uses water and tissue concentrations, diagnostic residues, and biological effects to set local criteria for hydrological units. This technique appears as Chapter 7 in my new book *Selenium Assessment in Aquatic Ecosystems* (Lemly 2002). I also present methods for setting environmentally safe ecosystem loading limits (TMDLs, Chapter 8) and delineating hydrological units (Chapter 6). Together, these 3 chapters can provide the guidance necessary to address site-specific questions. The book would appear to be essential for states and tribes in order for them to properly translate the new tissue-based criterion into state- or site-specific standards and limits on selenium sources/discharges. I recommend that, at a minimum, Lemly (2002) be added to the references cited at the end of the "Implementation" section (pg. 66), as well as to the document's reference list (pg. Ref-47). I am enclosing a copy of my book along with this review for EPA's information and consideration. If EPA endorses my procedure, it would be appropriate to formally mention it in the last paragraph on page 66 as the site-specific methodology recommended for use by states/tribes to set local water standards to meet the tissue criterion. I have supplied suggested wording in the margin. There is no hidden agenda here – I do not receive royalties from the sale of this book (my Federal employee status prevents it). I wrote it as a service to all those involved in selenium pollution issues, and it appears that it will have much more application now that a national tissue-based criterion is being proposed by EPA. My only interest is in making sure that those who need to conduct hazard assessments and set site-specific standards for selenium have proper guidance on how to go about it.

***Response*:** The comments are well taken. The reviewer's procedure for deriving site-specific chronic criteria for selenium appear geared toward modification of a water-column criteria concentration, such as EPA's 1987 criterion, but will nevertheless be considered during development of implementation guidance. As noted above, language was added to the revised draft document to allow for the development of a site-specific FCV where appropriate. EPA also plans to issue guidance to states and tribes on the implementation of the tissue-based criterion.

8. Although the criterion was not derived using wildlife criteria derivation procedures, EPA noted some evidence that the criterion would protect piscivorous birds. Are you aware of other data relevant to the protectiveness of the criterion for birds?
Comment: I am not aware of other data relevant to the protectiveness of the criterion for birds.

***Response*:** So noted. Nevertheless, in response to FWS concerns, EPA has agreed to remove all reference to birds.

Specific Comments:

Comment: 1. Since the new chronic criterion is tissue-based, the title of the document would be more accurate if the word tissue was added, i.e., Aquatic Life Water *and Tissue* Quality Criteria for Selenium. Similarly, on page 1, 2nd paragraph, 1st sentence ....water *and tissue* quality criteria....

***Response*:** We do not agree, although technically, considered outside our legal context, the comment has some merit. Nevertheless, such terminology is not appropriate within regulatory context of aquatic life water quality criteria and supporting programs.

Comment: 2. I have several comments written on the margin of pages 53-57. These are intended to improve the validity of the discussion, not change the EC_{20s} or chronic values.

Part A. ...The statement on page 53, 3rd paragraph, "In some field studies, chronic tolerance to selenium appears to be much higher than in laboratory studies", is misleading and does not correctly interpret the literature cited. The underlying problem is that the authors are improperly equating different exposure levels, life stages, and biological effects endpoints in order to support this assertion. It is important to know that data for the polluted areas of Belews Lake and Hyco Reservoir come from ecosystems that were saturated with selenium, including high residues in fish. It is not surprising, therefore, that the resulting chronic values and EC_{20s} based on these data are much higher than for controlled laboratory and stream studies where exposures (food chain residues) were manipulated and were limited to low-to-moderate concentrations. This is why the lab and stream studies were done in the first place, that is, to ascertain the low-level and threshold effects, since the field food chains were grossly polluted and preliminary evidence suggested severe impacts on fish. Concentrations in field biota were typically well above the highest levels of contamination observed in the lab/stream – not surprisingly, samples of tissue yielded relatively high chronic values and $EC_{20s.}$ There were no "low selenium" field exposures to provide an estimate of the threshold response or to bracket the lowest effect level as there were for the controlled studies. If you only have high levels of contamination to sample, you're only

going to get high tissue residues, chronic values, and EC_{20s}. What the authors suggest as tolerance in natural field settings can be explained simply on the basis of different exposure conditions.

Part B. ...Another problem concerns the authors' inappropriate mixing of endpoints. For example, the EC_{20} for deformities in Belews Lake fish, and the chronic value for the Bryson and Gillespie (BG) field studies, are not directly comparable to the effect thresholds determined in the Hernmanutz and Coyle (HC) studies. For one thing, the HC number represents an effect threshold whereas the chronic values for the BG studies are a "concentration in the female parent associated with this high occurrence of mortality of hatched larvae" (page 53, 4th paragraph), clearly not a threshold, and the Belews number is a 20% effect level, also not a threshold – the authors incorrectly compare three different effect levels. No wonder the HC threshold is "approximately 3 times lower than those recorded above" (page 54, 2nd paragraph). Also, the Belews number is for teratogenic deformities, the BG number is for parent fish, but the HC studies used larva/fry survival – the authors mix three different endpoints. It is not valid to compare different effect levels and endpoints, incorrectly equate them, and say (speculate) that the difference is due to tolerance. Moreover, the Belews number is generated from data for juvenile and adult fish, which automatically yields a two-fold higher value than for larvae and fry (see Lemly 2002, page 94 for a figure that explains this difference). Thus, the life stage of fish must be accounted for when evaluating teratogenic deformity data – yet another endpoint consideration that the authors failed to recognize. With such a mixing of effect levels and biological endpoints, it is not surprising that the chronic value is higher for field studies, but this does not indicate tolerance. The authors have presented an analysis that seems plausible on the surface, but it is not technically valid if one examines the data sets, endpoints, and associated effect levels carefully. There are 4 major points that should be addressed in order to provide a valid interpretation of the data for the discussion on pages 53-58: (1) Don't compare tissue residues from fish with different levels of exposure (i.e., 30% or greater difference in environmental concentrations and/or dietary intake) and then try to infer that differences in the resultant chronic values or EC_{20s} are due to tolerance – only compare same-level exposures, (2) Don't mix threshold values with high mortality impacts or EC_{20s} – these are not the same effect levels, (3) Don't mix the endpoints – keep fry survival, parental concentrations, and teratogenic effects separate, and (4) When evaluating teratogenic effects, don't mix life stages – keep larvae/fry separate from juveniles/adults.

Part C. ...The authors need to make the appropriate data comparisons and then modify the discussion. The suggestion of tolerance is simply not valid as they present it, and should be dropped altogether. However, on page 58, the decision that Bryson and Lemly not be used in the bluegill GMCV is valid. The authors exclude these studies because they (authors) infer tolerance developed from multiple-generations of exposure when, in fact, the reason is different levels of exposure across generations (i.e., declining selenium in the food chain over time coupled with recolonization and artificial stocking of fish; see Lemly 1997a for data and discussion; also Bryson 1985a, pg. 2-9, paragraph 2), not generational differences in sensitivity. The authors arrived at the right conclusion, but for the wrong reason. Moreover, the 19 µg/g dw "no effect" level that the criterion document gives (pg. 56, last sentence) for bluegill larvae in Bryson's study is not reliable because: (1) There were no replicates for the tissue analyzed, and thus "sample

sizes were not sufficient for statistical comparison of concentrations" (pg. 2-8 in Bryson) – i.e., no geometric mean can be determined, and (2) The trend in selenium concentrations relative to percent effluent in the nonaffected area was reversed compared to that for both the control (Roxboro City Lake) and the affected area (Table 2.8, pg. 2-17 in Bryson), that is, as percent effluent increased so did selenium concentrations in the control and affected areas, but the relationship was opposite in the nonaffected area (from which the 19 µg/g value is taken). This apparent discrepancy may have been due to influences of mercury on selenium bioaccumulation and toxicity (Bryson, pg.2-9, paragraph 3), but it remains an unanswered question that clouds the interpretation of data from the "unaffected" portion of Hyco Reservoir. It is incorrect for the authors to infer that Bryson's 19 µg/g dw value is indicative of tolerance – this statement (last sentence on page 56) needs to be removed from the criterion document, as well as the last sentence on page 58.

Response: We agree that for most of the field studies the LOAEC values are relatively high because of the high levels of selenium in the system. However, in addition to the Kennedy et al (2000) study, the Bryson et al (1985a) study reported no effects to swim-up larvae that came from adults containing 19.18 µg/g dw. We also agree the endpoint of teratogenic deformities is not comparable to embryo-larval effects. The Lemly (1993a) reference is not be used to compare the sensitivity between laboratory and field exposed organisms in the revised draft document.

We do not agree with the statements cited by the reviewer as reasons why the 19.18 µg/g dw tissue value is not reliable. Table 2.7 from Bryson et al (1985a) indicates selenium was measured from 2 replicate females from the unaffected area in Hyco Reservoir (average whole body selenium concentration equal to 19.18 µg/g dw) that produced the swim-up larvae which did not show deleterious effects. The reviewer stated that the selenium measurements were not replicated. The reviewer also suggested that the occurrence of the inverse concentration relationship in the larvae from the unaffected area exposed to 0, 20 and 50 percent effluent (Table 2.8 in Bryson et al.) adds to the unreliability of the 19.18 µg/g dw value. The selenium concentrations listed in Table 2.8 in Bryson et al (1985a) are for surviving bluegill larvae from the embryo-larval experiments; they are not representative of parent tissue concentrations.

As it is written now, the sentence... "In some field studies, chronic tolerance to selenium appears to be much higher than in laboratory studies..." does imply that tolerance to selenium is a real phenomenon exhibited by wild fish populations living in highly contaminated sites, though in fact, this ***hypothesis*** has not been rigorously tested (see Kennedy et al. 2000). The original statement was made as a possible explanation for the differences in chronic toxicity observed between fish exposed to selenium naturally in the field and those examined in the laboratory. Further study is required to either reject or accept this hypothesis. In order to dispel any concern of the misleading statement made on page 53, paragraph 3, the text was changed on the revised text (page 73) to read:

> "It should be noted that the acquisition of tolerance to selenium has been hypothesized (Kennedy et al. 2000), but has not yet been substantiated."

Comment: 3. Page 58, 3rd paragraph, states "This appeared to reduce the tissue concentration associated with reduced survival". Lemly 1993a, Fig. 1, shows an increase in selenium in the group that exhibited reduced survival (shown as cold + Se, Fig. 9). This sentence needs to be reworded to something like "Cold water temperature increased the sensitivity of fish to selenium, but appeared to increase the tissue concentration associated with reduced survival (5-6 µg/g dw @ 20° C versus 7-8 µg/g dw @ 4° C)". I use the word "appeared" because in the discussion of Lemly 1993a, he attributes the difference in residues to the reduced lipid content of the affected group (lipid has low Se relative to other body constituents), which served to increase the concentration in remaining whole-body tissues. Thus, a change in body chemistry, rather than a change in total body burden of selenium, was likely responsible for the "apparent" increase in tissue concentrations.

***Response*:** So noted. The text in the revised draft document has been changed considerably and reflects the above suggestions.

Response to Additional Unsolicited Comments from

Dennis Lemly
U.S. Forest Service, Southern Research Station
Coldwater Fisheries Research Unit
1650 Ramble Road, Blacksburg, VA 24060

June 22, 2004

EPA received the following comment in a letter from Dennis Lemly to Denise Keehner two years after completion of the peer review. EPA has chosen to include it as part of the peer review.

Comment: I am writing to formally retract my prior peer review of EPA's draft "Aquatic Life Water Quality Criteria for Selenium – 2002", as transmitted to you via my letter dated April 23, 2002. Since that time, several issues have been brought to my attention which raise serious questions as to the scientific credibility of the draft document. Chief among the concems are: (1) the full utility of the experimental design of my 1993 winter stress study, which was the controlling experiment used by EPA's contractor (Great Lakes Environmental Center) to derive the 7.9 µg/g dw chronic criterion, went unrecognized, (2) the crucial linear regression equation relating selenium concentrations in fish ovaries to concentrations on a whole-body basis was erroneously reported, and (3) assessments of risk to aquatic-dependent wildlife were based on out-of-date information. I have re-examined my own study as well as original data and reports of other research cited in the draft document. In the course of this analysis it became clear that the 3 concerns listed above are indeed legitimate and, I believe, constitute fatal flaws in the calculations and reasoning used by GLEC to prepare the draft; ergo, the chronic criterion should be substantially lower than 7.9. A detailed discussion of these concerns has been provided to the EPA (Charlie Delos) on June 16, 2004 in the form of a draft manuscript written by several colleagues and myself. Because of the serious flaws I cannot ethically continue to let my previous endorsement of the 7.9 criterion stand. Please take notice that I no longer support the criterion as proposed in the draft document. I will be glad to discuss this matter further if you wish (540-231-6663; dlemly@fs.fed.us).

Response: Regarding the comment's Item 1, the interpretation of the Lemly 1993 cold temperature study, EPA has changed both its discussion of the study and its 2004 criteria statement.

(A) Mortality is now correctly described as 40% for the study's single treatment concentration. Readers should note that subsequent communications with the reviewer indicate that Lemly's 1993 publication contains an error. Each replicate began with 80 individuals, not 70 individuals.

Because the study has only a single treatment concentration (plus control), no concentration-response curve can be constructed and no EC20 or Chronic Value can be estimated by standard procedures.

(B) The tissue concentration at Day 60, 5.85 µg/g, is now part of the criterion. Although 5.85 µg/g is below the threshold for effects on bluegill at 20 °C, EPA agrees that under the conditions

of the study, the detection of 5.85 µg/g in summer or fall could be a harbinger of overwinter elevation of tissue concentrations to 7.91 µg/g, yielding mortality. Consequently, the 2004 criterion statement now says that if fish tissue concentrations exceed 5.85 µg/g dw during summer or fall, fish tissue should be monitored during the winter to determine whether the selenium concentration exceeds 7.91 µg/g.

In requesting information, data, and views from the public, EPA specifically flagged the interpretation and use of the Lemly 1993 study as an issue of concern. EPA will be revisiting these matters.

Regarding Item 2, reviewer Greg Moller noted the transcription error in the ovary to whole body conversion. Moller also correctly recognized that the transcription error only affected the final presentation of the equation and not the original calculations.

Applied to the ovary equation presentation error, EPA considers Lemly's appellation "fatal flaw" to be a hyperpole. In any case, all the regressions have been redone in response to other comments of reviewer Greg Moller and other communications from Dennis Lemly and others.

Regarding Item 3, EPA has removed all discussion of wildlife from the document, per an agreement with FWS.