

EPA RESPONSE TO EXTERNAL PEER REVIEWER COMMENTS
on
DRAFT 2009 UPDATE AQUATIC LIFE
AMBIENT WATER QUALITY CRITERIA FOR
AMMONIA - FRESHWATER

April 2013

Office of Water
U.S. Environmental Protection Agency
Washington, DC

EPA RESPONSE TO EXTERNAL PEER REVIEWER COMMENTS

on

**DRAFT 2009 UPDATE AQUATIC LIFE
AMBIENT WATER QUALITY CRITERIA FOR
AMMONIA - FRESHWATER**

Prepared for:

U.S. Environmental Protection Agency
Health and Ecological Criteria Division
Office of Science and Technology
Office of Water
Washington, DC

Prepared by

Great Lakes Environmental Center
Columbus, Ohio

April 2013

TABLE OF CONTENTS

1.	Original Charge Questions.....	1
1.1.	Background.....	1
1.2.	Document that Was Reviewed.....	1
1.3.	Charge Questions.....	1
1.4.	Reviewers.....	2
2.	Acute criteria in fresh waters.....	3
2.1.	Question 1: Are the toxicity tests used to derive the criteria scientifically defensible for such use? Are you aware of other relevant data that were not used?.....	3
2.1.1.	Peer Reviewer 1.....	3
2.1.2.	Peer Reviewer 2.....	3
2.1.3.	Peer Reviewer 3.....	4
2.1.4.	Peer Reviewer 4.....	5
2.1.5.	Peer Reviewer 5.....	5
2.1.6.	Peer Reviewer 6.....	5
2.1.7.	Peer Reviewer 7.....	6
2.2.	Question 2: What are the technical considerations that EPA should evaluate when mussels are present and mussels are absent with respect to the recommended acute criteria?.....	7
2.2.1.	Peer Reviewer 1.....	7
2.2.2.	Peer Reviewer 2.....	9
2.2.3.	Peer Reviewer 3.....	9
2.2.4.	Peer Reviewer 4.....	10
2.2.5.	Peer Reviewer 5.....	10
2.2.6.	Peer Reviewer 6.....	11
2.2.7.	Peer Reviewer 7.....	11
2.3.	Question 3: Is it scientifically defensible to exclude the glochidia data at this time due to the uncertainty of appropriate test duration time for this life stage? Do you believe there is an alternative approach to the use of this data that would be more scientifically sound?.....	12
2.3.1.	Peer Reviewer 1.....	12
2.3.2.	Peer Reviewer 2.....	13
2.3.3.	Peer Reviewer 3.....	14
2.3.4.	Peer Reviewer 4.....	15
2.3.5.	Peer Reviewer 5.....	15
2.3.6.	Peer Reviewer 6.....	15
2.3.7.	Peer Reviewer 7.....	16
2.4.	Question 4: Regarding the proposed approach to glochidia data in the 2009 draft position statement as it relates to ecological relevance and practicality - Is the approach a scientifically defensible principle for structuring the population exposure duration problem and designing further research to quantify such a duration?.....	16
2.4.1.	Peer Reviewer 1.....	16
2.4.2.	Peer Reviewer 2.....	17
2.4.3.	Peer Reviewer 3.....	18
2.4.4.	Peer Reviewer 4.....	18
2.4.5.	Peer Reviewer 5.....	18
2.4.6.	Peer Reviewer 6.....	18
2.4.7.	Peer Reviewer 7.....	18

3.	Hyaella azteca position statement and proposed rationale:	19
3.1.	Question 5: Are the position statement and supporting rationale regarding use of toxicity data for <i>Hyaella azteca</i> in criteria development reasonable and scientifically sound recommendations?	19
3.1.1.	Peer Reviewer 1	19
3.1.2.	Peer Reviewer 2	21
3.1.3.	Peer Reviewer 3	21
3.1.4.	Peer Reviewer 4	21
3.1.5.	Peer Reviewer 5	21
3.1.6.	Peer Reviewer 6	22
3.1.7.	Peer Reviewer 7	22
4.	Chronic Freshwater Criterion:	22
4.1.	Question 6: Are the toxicity tests and other studies used to derive the criterion scientifically defensible for such use? Are you aware of other relevant data that were not used?.....	22
4.1.1.	Peer Reviewer 1	22
4.1.2.	Peer Reviewer 2	23
4.1.3.	Peer Reviewer 3	25
4.1.4.	Peer Reviewer 4	25
4.1.5.	Peer Reviewer 5	25
4.1.6.	Peer Reviewer 6	26
4.1.7.	Peer Reviewer 7	26
4.2.	Question 7: Is the freshwater chronic criterion scientifically defensible with mussels present and mussels absent?.....	26
4.2.1.	Peer Reviewer 1	26
4.2.2.	Peer Reviewer 2	27
4.2.3.	Peer Reviewer 3	27
4.2.4.	Peer Reviewer 4	28
4.2.5.	Peer Reviewer 5	28
4.2.6.	Peer Reviewer 6	28
4.2.7.	Peer Reviewer 7	28
5.	Use of 28-day Juvenile Test Data:	29
5.1.	Question 8: Given that the juvenile life stage of freshwater mussels is relatively long (2-6 years) are 28-day exposure tests with juvenile mussels scientifically defensible as "chronic" test data for criteria development?.....	29
5.1.1.	Peer Reviewer 1	29
5.1.2.	Peer Reviewer 2	30
5.1.3.	Peer Reviewer 3	31
5.1.4.	Peer Reviewer 4	32
5.1.5.	Peer Reviewer 5	32
5.1.6.	Peer Reviewer 6	32
5.1.7.	Peer Reviewer 7	32
5.2.	Question 9: Should toxicity studies on the growth rates of mussel shells during 28-day tests be considered quantitatively when developing water quality criteria?.....	33
5.2.1.	Peer Reviewer 1	33
5.2.2.	Peer Reviewer 2	33
5.2.3.	Peer Reviewer 3	34
5.2.4.	Peer Reviewer 4	34
5.2.5.	Peer Reviewer 5	34
5.2.6.	Peer Reviewer 6	34
5.2.7.	Peer Reviewer 7	35

5.3. Question 10a: Regarding the position statement and rationale on use of juvenile mussel growth data – Is it scientifically defensible to include the juvenile growth data from a 28-day exposure period as “other data” that might influence the criteria however not be used directly in the derivation of the criteria value?.....	35
5.3.1. Peer Reviewer 1	35
5.3.2. Peer Reviewer 2	35
5.3.3. Peer Reviewer 3	36
5.3.4. Peer Reviewer 4	36
5.3.5. Peer Reviewer 5	36
5.3.6. Peer Reviewer 6	36
5.3.7. Peer Reviewer 7	36
5.4. Question 10b: Regarding the position statement and rationale on use of juvenile mussel growth data –Should the statement also consider impaired growth of mussels which were affected at a 28-day exposure could as likely continue to decline in longer exposures as another potential outcome (i.e., the chance they could recover or stabilize is one potential outcome only)?.....	37
5.4.1. Peer Reviewer 1	37
5.4.2. Peer Reviewer 2	37
5.4.3. Peer Reviewer 3	37
5.4.4. Peer Reviewer 4	37
5.4.5. Peer Reviewer 5	37
5.4.6. Peer Reviewer 6	38
5.4.7. Peer Reviewer 7	38
5.5. Question 11: The values of the acute and chronic ammonia criteria have a strong dependence on pH. Juvenile and adult mussels, as sediment-dwelling organisms, inhabit a medium that may have vertical pH gradients, thereby creating some uncertainty about the appropriate pH to assign as their exposure conditions. For applying a criterion protecting mussels, do you have suggestions on how states and EPA might determine the pH applicable to the sediment micro-environment to which mussels are typically exposed?	38
5.5.1. Peer Reviewer 1	38
5.5.2. Peer Reviewer 2	39
5.5.3. Peer Reviewer 3	39
5.5.4. Peer Reviewer 4	40
5.5.5. Peer Reviewer 5	40
5.5.6. Peer Reviewer 6	40
5.5.7. Peer Reviewer 7	41
5.6. Question 12: In general, should the criteria include a consideration for the potential pH difference between sediment and the water? If so, what is the most scientifically defensible way to account for these differences when deriving protective water quality criteria?.....	41
5.6.1. Peer Reviewer 1	41
5.6.2. Peer Reviewer 2	41
5.6.3. Peer Reviewer 3	42
5.6.4. Peer Reviewer 4	42
5.6.5. Peer Reviewer 5	42
5.6.6. Peer Reviewer 6	42
5.6.7. Peer Reviewer 7	42
5.7. Question 13: Should exposure tests on juvenile mussels be conducted with or without sediment in the test chamber?	43
5.7.1. Peer Reviewer 1	43
5.7.2. Peer Reviewer 2	45

5.7.3.	Peer Reviewer 3	45
5.7.4.	Peer Reviewer 4	46
5.7.5.	Peer Reviewer 5	46
5.7.6.	Peer Reviewer 6	46
5.7.7.	Peer Reviewer 7	46
5.8.	Additional Comments	47
5.8.1.	Peer Reviewer 3	47
5.8.2.	Peer Reviewer 4	49
5.8.3.	Peer Reviewer 6	49

1. ORIGINAL CHARGE QUESTIONS

The document reviewed was a draft update of EPA's 1999 aquatic life criterion for ammonia in freshwater. The draft document was titled: Draft Reassessment of the 1999 Ambient Water Quality Criteria for Ammonia – Freshwater, dated July 24, 2009.

1.1. Background

The aquatic life criterion for ammonia is used for determining the level of pollution control necessary to attain the aquatic life uses of water bodies, and in particular to allow the propagation of fish and shellfish, per Section 302 of the Clean Water Act.

In a 2004 interagency memo EPA stated that it would conduct a criteria re-evaluation of the 1999 ammonia aquatic life criteria in light of new studies indicating that freshwater mussels may be more sensitive to ammonia than the organisms considered in the 1999 Update of Ambient Water Quality Criteria for Ammonia. In addition, a Federal Register Notice (FRN) was published in 2004 notifying the public of the criteria re-evaluation and requesting any pertinent toxicity data. Responding to comments from the FRN regarding concerns about uncertainty in the new mussel test protocol and data, EPA hosted the Mussel Toxicity Testing Protocol workshop in 2005 to convene experts to discuss the new test protocol for glochidia and juvenile mussels.

The 2009 draft criteria update of the 1999 Update incorporated new toxicity data on freshwater mussels as well as snails and some ESA listed fish species.

1.2. Document that Was Reviewed

The document reviewed was titled "Reassessment of the 1999 Ambient Water Quality Criteria for Ammonia - Freshwater". The document presented acute and chronic ammonia criteria for fresh waters with mussels present and a separate criterion for waters with mussels absent.

1.3. Charge Questions

Acute criteria in fresh waters:

1. Are the toxicity tests used to derive the criteria scientifically defensible for such use? Are you aware of other relevant data that were not used?
2. What are the technical considerations that EPA should evaluate when mussels are present and mussels are absent with respect to the recommended acute criteria?
3. Is it scientifically defensible to exclude the glochidia data at this time due to the uncertainty of appropriate test duration time for this life stage? Do you believe there is an alternative approach to the use of this data that would be more scientifically sound?
4. Regarding the proposed approach to glochidia data in the 2009 draft position statement as it relates to ecological relevance and practicality - Is the approach a scientifically defensible principle for structuring the population exposure duration problem and designing further research to quantify such a duration?

Hyalella azteca position statement and proposed rationale:

The EPA workgroup developed a position statement and proposed supporting rationale describing the concerns over using *Hyalella azteca* toxicity test data in criteria development due to the uncertain health of the test organisms in different test water composition. The rationale defined the specific concerns and

uncertainties supporting the recommended exclusion of the *Hyaletta* data from use in criteria derivation, at that time. The position statement was based on the workgroup's review of a number of toxicity tests on *Hyaletta*, referenced in the rationale.

5. Are the position statement and supporting rationale regarding use of toxicity data for *Hyaletta azteca* in criteria development reasonable and scientifically sound recommendations?

Chronic Freshwater Criterion:

6. Are the toxicity tests and other studies used to derive the criterion scientifically defensible for such use? Are you aware of other relevant data that were not used?
7. Is the freshwater chronic criterion scientifically defensible with mussels present and mussels absent?

Use of 28-day Juvenile Test Data:

Water quality criteria for the protection of aquatic life are derived using toxicity endpoints that relate to population level impacts. In general, these endpoints relate to survival, growth and/or reproduction. The 28-day test with juvenile mussels, while similar in duration to a standard chronic test, is not technically an early-life stage test according to the 1985 Guidelines for Aquatic Life Criteria as much of the early development will have already occurred.

8. Given that the juvenile life stage of freshwater mussels is relatively long (2-6 years) are 28-day exposure tests with juvenile mussels scientifically defensible as "chronic" test data for criteria development?
9. Should toxicity studies on the growth rates of mussel shells during 28-day tests be considered quantitatively when developing water quality criteria?
10. Regarding the position statement and rationale on use of juvenile mussel growth data:
 - 10a. Is it scientifically defensible to include the juvenile growth data from a 28-day exposure period as "other data" that might influence the criteria however not be used directly in the derivation of the criteria value?
 - 10b. Should the statement also consider impaired growth of mussels which were affected at a 28-day exposure could as likely continue to decline in longer exposures as another potential outcome (i.e., the chance they could recover or stabilize is one potential outcome only)?
11. The values of the acute and chronic ammonia criteria have a strong dependence on pH. Juvenile and adult mussels, as sediment-dwelling organisms, inhabit a medium that may have vertical pH gradients, thereby creating some uncertainty about the appropriate pH to assign as their exposure conditions. For applying a criterion protecting mussels, do you have suggestions on how states and EPA might determine the pH applicable to the sediment micro-environment to which mussels are typically exposed?
12. In general, should the criteria include a consideration for the potential pH difference between sediment and the water? If so, what is the most scientifically defensible way to account for these differences when deriving protective water quality criteria?
13. Should exposure tests on juvenile mussels be conducted with or without sediment in the test chamber?

1.4. Reviewers

In July of 2009 the Charge Questions along with a copy of EPA's *DRAFT REASSESSMENT OF THE 1999 AMBIENT WATER QUALITY CRITERIA FOR AMMONIA – FRESHWATER*, dated July 24, 2009

(the title of which subsequently was changed to *DRAFT 2009 UPDATE AQUATIC LIFE AMBIENT WATER QUALITY CRITERIA FOR AMMONIA – FRESHWATER*, EPA-822-D-09-001 for publication in the Federal Register), was sent to seven reviewers and included: Alex Barron, Steve P. Canton, M.S., Jerome M. Diamond, Ph.D., George Dixon, Ph.D., Jerry L. Farris, Ph.D., William L. Goodfellow Jr., M.S., and James J. Pletl, Ph.D. The sections below are organized by each Charge Question with peer reviewers' [verbatim] comments and EPA response for each charge question.

2. ACUTE CRITERIA IN FRESH WATERS

2.1. Question 1: Are the toxicity tests used to derive the criteria scientifically defensible for such use? Are you aware of other relevant data that were not used?

2.1.1. Peer Reviewer 1

Yes, all the toxicity tests that influence the derivation of the new criteria for ammonia appear to be scientifically defensible. All relevant data that I am aware of has been considered in this reassessment document.

2.1.1.1. EPA Response

Thank you for your comment.

2.1.2. Peer Reviewer 2

Yes, in general the toxicity tests used to derive the water column criteria, as cited in the draft update, are scientifically defensible. Plus, it appears they provide sufficient data to evaluate inter and intra-species comparisons using the guidelines established for deriving water quality criteria.

Additionally, a few years ago, we conducted a review of a number of ambient water quality criteria (AWQC) as part of the Arid West Water Quality Research Project (AWWQRP 2006). One aspect of that project included a review and update to the 1999 EPA Ammonia criteria document, similar to this draft. Checking those updates side-by-side, it appears there are a few more studies that could be included in this re-evaluation:

- 1) Augspurger, T., A.E. Keller, M.C. Black, W.C. Cope, and F.J. Dwyer. 2003. Water quality guidance for protection of freshwater mussels (Unionidae) from ammonia exposure. *Environmental Toxicology and Chemistry* 22:2569-2575.
- 2) Khatami, S. H., D. Pascoe, M. A. Learner. 1998. The acute toxicity of phenol and unionized ammonia, separately and together, to the ephemeropteran *Baetis rhodani* (Pictet). *Environmental Pollution* 99:379-387.
- 3) Rubin, A.J., and M.A. Elmaraghy. 1977. Studies on the toxicity of ammonia, nitrate and their mixtures to guppy fry. *Water Research* 11:927-935.
- 4) Russo, R.C., D.J. Randall, and R.V. Thurston. 1988. Ammonia toxicity and metabolism in fishes. In R.C. Ryans, ed. *Protection of River Basins, Lakes, and Estuaries*. pp. 159-173. American Fisheries Society, Bethesda, Maryland.
- 5) Sangli, A. B. and V. V. Kanabur. 2001. Toxicity of ammonia to a freshwater fish, *Gambusia affinis* and its effect on oxygen consumption. *Geobios* 28:56-58.
- 6) Tomasso, J.R., C.A. Goudie, B.A. Simco, and K.B. Davis. 1980. Effects of environmental pH and calcium on ammonia toxicity in channel catfish. *Transactions of the American Fisheries Society* 109:229-234.

I should also note that in that prior review of the 1999 Ammonia document (AWWQRP 2006), we raised concerns regarding the appropriateness of some studies/data used in the 1999 EPA document (see Table 1

below). These studies may need an additional evaluation to determine if they provide toxicity data meeting EPA data quality guidelines (Stephan et al. 1985).

Deletions of Inappropriate Data Presented in the EPA 1999 AWQC Ammonia Document (from AWWQRP 2006)

Species	References	Comments
Acute		
<i>Simocephalus vetulus</i>	Mount 1982	Insufficient data to validate results
<i>Oncorhynchus mykiss</i>	Calamari et al. 1977	Unable to validate citation or data-other data available by same author with this species
<i>Oncorhynchus mykiss</i>	Thurston et al. 1981	Unable to validate citation or data
<i>Oncorhynchus mykiss</i>	Reinbold & Pescitelli 1982	Unable to validate citation or data
<i>Pimephales promelas</i>	Thurston et al. 1981	Unable to validate citation or data
<i>Pimephales promelas</i>	Reinbold & Pescitelli 1982	Unable to validate citation or data
<i>Gambusia affinis</i>	Wallen et al. 1957	Insufficient data to validate results
<i>Lepomis macrochirus</i>	Reinbold & Pescitelli 1982	Unable to validate citation or data
<i>Sander vitreum</i>	Reinbold & Pescitelli 1982	Unable to validate citation or data
Chronic		
<i>Ictalurus punctatus</i>	Colt & Tchobanoglous 1978	Insufficient data to validate results

2.1.2.1. EPA Response

*EPA reviewed the list of suggested references and several (#1, 3, 4, 6 as listed above) were used as acceptable data in the final ammonia criteria document. Reference #2 was not used because it did not meet 1985 Guidelines requirements for duration of test and residency in North America (see App. L p. 186 in 2013 ammonia document). Reference #5 was not included as acceptable data in the final ammonia criteria because identified the test species (*Lepidocephalichthys guntea*) was also as a freshwater fish species not resident in North America.*

The list of references that are suggested for deletion in the above table were reviewed and concluded to be acceptable data in the final ammonia criteria document. All of the above citations are published in the scientific literature and are publicly available.

2.1.3. Peer Reviewer 3

In general, the toxicity tests used to derive ammonia criteria in the draft Reassessment Document (RD) are scientifically defensible given EPA’s criteria development guidelines (Stephan et al. 1985). However, I have reservations regarding the use of snail data derived from field-collected organisms as explained below under Chronic Freshwater Criterion. I am not aware of other relevant freshwater toxicity test data that were not either used or considered in the RD.

2.1.3.1. EPA Response

Thank you for your comment.

Note: Upon subsequent review of the repeat snail test by Besser (2011) – see bibliographic citation provided in the 2013 ammonia criteria document - this reviewer agreed the data for the larger sized snails in the test was scientifically defensible for use in the derivation of the chronic ammonia criterion.

2.1.4. Peer Reviewer 4

I am familiar with the majority of the literature cited and, in my opinion, the tests used to develop the acute criteria are scientifically defensible. I am, however, relying on the fact that the Draft Reassessment states that a process to identify data meeting the minimum requirements for inclusion was followed. I have not obtained all of the cited literature and reviewed each paper for suitability of inclusion of the findings in the data sets; this type of detailed audit is beyond the scope of this review. I follow the ammonia literature closely, and I am unaware of any additional data that could be of use.

2.1.4.1. EPA Response

Thank you for your comment.

2.1.5. Peer Reviewer 5

With greater recent attention dedicated to standardization of test techniques suitably addressing the feasibility of repeatability and precision of results from juvenile mussel responses (Augspurger, 2007), data available at the time of this reassessment meet a higher standard of acceptability for guidance modified from the US EPA (Stephan et al., 1985) and follow consensus test protocols and quality assurance recommendations sufficient to support water quality criteria development. The majority of tests used to derive the criteria seem scientifically defensible for such use. Upon review of those included and others not specifically included in the calculation, but cited in Table 1 of the draft addendum, there are tests that present concern for acclimation procedures of test organisms and culturing (either reported or unreported) and lack of reference testing that can insure level of expertise or consistency associated with such field collection, organism transfer, and culturing that can significantly influence test responses. Specifically, sufficient acclimation information for water quality and temperature was not included in Mummert et al., 2003 and Scheller, 1997.

I am not aware of other data from standardized toxicity tests with ammonia that could be considered relevant to this reassessment.

2.1.5.1. EPA Response

Acclimation information for water quality and temperature is included in Mummert et al. (2003) on page 2546 of the article (Materials and Methods under Juvenile mussel bioassays). Methods used are in accord with established laboratory toxicity testing guidance for the taxa (see ASTM E 2455-06).

While it is true that information specific for water quality and temperature acclimation is generally lacking in the Scheller (1997) Master's Thesis, sufficient information is provided on type of test dilution water used (i.e., filtered and aerated Sinking Creek water from Newport, VA; a previously well described and characterized and test dilution water used to support aquatic toxicity testing at Virginia Polytechnic Institute in Blacksburg, VA) and test conditions such that EPA has no reason to exclude these data. In addition, the toxicity test results from the study are consistent with other results for the species and survival of control animals was acceptable and/or were or were not included in the criteria's numeric formulation.

(Note: See bibliographic citations provided in the 2013 ammonia criteria document for these and all subsequent references noted in responses, unless indicated otherwise.)

2.1.6. Peer Reviewer 6

From my review, I believe that the acute toxicity testing data is relevant. My only concern is that this document was careful to use data that was from published references which I believe is a strength of this document as compared to earlier water quality criteria documents. Since the only unpublished data used that demonstrates a somewhat low toxicity response to ammonia was from Keller 1999 for the pondshell

mussel, maybe putting a synopsis as an attachment would be warranted. This would help make the document transparent as the other Attachments help do.

The Table A in text of the GMAV for the mountain whitefish does not match Table 1. The Text table reports 12.11 and Table 1 reports 12.09. I am expecting that the Table 1 is correct and is what was used for the FAV. Please verify.

2.1.6.1. EPA Response

The Keller 1999 memo is available in the docket. The GMAV values for Mountain whitefish have been corrected: the GMAV at pH 8 is 12.11 mg TAN/L, and the GMAV at pH 7 is 51.93 mg TAN/L.

2.1.7. Peer Reviewer 7

The question requires that the term “scientifically defensible” be defined within the context of the use or the consequences of the use. The use is to support development of water quality criteria that will be promulgated by states as legally enforceable water quality standards. Failure to meet water quality standards instream and/or at the edge of respective mixing zones will result in TMDLs, implementation plans, or NPDES permit limits for dischargers to the respective stream. Failure to meet water quality standards also infers that designated uses are not being supported. In the case of ammonia criteria, the designated use is aquatic life. The Clean Water Act also states that violations of such permit limits can result in significant monetary fines and/or jail time for responsible individuals. Therefore the use, in this case, is quite demanding of the data.

The level of scientific defensibility set by EPA for this use is limited to that defined by EPA’s 1985 “Guidelines For Deriving Numerical National Water Quality Criteria For The Protection of Aquatic Organisms And Their Uses”, hereafter referred to as the Guidelines, and the published, peer reviewed toxicity test methods used in each case. The tests in question appear to meet these standards; however the data quality objectives set by these documents may not be sufficient for the data use. The toxicity test methods and the Guidelines do not require what are considered to be critical elements of scientifically defensible toxicity tests. Acute tests typically only have one quality objective: minimum survival in controls. Other than control survival there are no other quality objectives that must be met for the results of these tests to be considered reliable. For example, 1) the methods and Guidelines do not require a concentration-response curve for data to be deemed acceptable for any use. There can be little certainty in a toxicity test result where response does not change predictably with the concentration exposure. If demonstration of this curve is not required of the test or EPA, data of questionable quality may be used to derive water quality criteria. 2) Analysis of acute toxicity test data typically also does not address the intra-treatment variability that commonly occurs in these tests; providing results with biased estimates on uncertainty. 3) The Guidelines also do not require replication of any test result for the most sensitive species. This would allow a single test result to significantly bias the calculation of water quality criteria. ASTM E2455-06 for testing freshwater mussels recommends that “different batches of the same species and the same life stage should be collected and tested over time in order to obtain a measure of the variability associated with testing the particular species”. This goal was not accomplished for the pink mucket, the third most acutely sensitive genus of the database.

Aside from universal issues for acute tests 4) there is also concern that the acute juvenile mussel tests may be biased towards findings of more sensitivity and toxicity. The tests proposed to be used for this reassessment are conducted without sediment in the test vessels. Newton and Bartsch (2007) found that without sediment juvenile organisms did not grow in 96 hour toxicity tests. This clearly shows that the organisms are stressed by the lack of sediment in the test vessels. This also strongly suggests that the survival endpoints from acute tests may be biased low (falsely indicating greater toxicity) due to the added stress of failing to provide sediment in the test vessels. All of these issues characterize a significant level of uncertainty which is indirectly proportional to scientific defensibility. EPA should address these uncertainties before using new data to update this criteria document.

I am not aware of other relevant data that were not used for this reassessment.

2.1.7.1. EPA Response

EPA relies on the best available scientific information to develop CWA Section 304(a) ambient water quality criteria for the protection of aquatic life. Studies from the open literature are reviewed against data quality objectives as described in the 1985 Guidelines in general, and other appropriate toxicity test guidance such as ASTM or OECD (Organization of Economic Co-operation and Development), specifically. When using the peer reviewed literature, it is not always possible to obtain the raw data for a particular study. For studies that are drivers of the criteria and/or are of critical importance to the assessment, EPA thoroughly analyzes the data, contacts researchers as necessary for additional information, and may even conduct an independent external peer review of the study(ies) as needed.

The Newton and Bartsch (2007) ammonia toxicity study cited above demonstrated that juvenile mussels (plain pocketbook, Lampsilis cardium) had reduced growth in tests without sediment and that lethal concentrations (LC50s) for those exposed in water (without sediment) were about two- to three-fold higher than those of juveniles exposed in sediment. This apparent difference in ammonia toxicity between exposure systems with and without sediment is thought to be a result of a pH difference between overlying water and porewater (see Wang et al. 2011; Environ. Toxicol. Chem. 30:2270). Moreover, it has since been shown that juvenile and adult mussels can survive and grow in culture without sediment at rates similar to those observed in the field if adequate water flow and suspended food are provided (Chris Barnhart, unpublished data; Wang et al. 2011).

Per EPA's internal supplement to the 1985 Guidelines (p. 11, #9 in Manual of Instructions for Preparing Aquatic Life Water Quality Criteria Documents, 12 December 1985), results from tests with sediment should generally not be used quantitatively in criteria development because such tests typically do not control for potential confounding factors and concentrations of a pollutant in the overlying test solution are not always measured. Currently the guidance for conducting toxicity tests with unionid mussels (ASTM E 2455-06) does not require acute toxicity tests to be conducted with sediment, and only considers survival of control animals at the end of the test for determining test acceptability.

All of the acute toxicity data and most of the chronic toxicity data for unionid mussels included in the final 2013 ammonia criteria are derived from water-only tests (without substrate). These studies represent the best available scientific data at this time and have been deemed acceptable for criteria development purposes at this time. [See "Aquatic Life Ambient Water Quality Criteria for Ammonia - Freshwater (2013)" page 35 - Summaries of studies used in chronic criterion determination: Lampsilis siliquoidea (fatmucket)].

2.2. Question 2: What are the technical considerations that EPA should evaluate when mussels are present and mussels are absent with respect to the recommended acute criteria?

2.2.1. Peer Reviewer 1

Table B compares Final Acute Value (FAV) and Criterion Maximum Concentration (CMC) values calculated with the entire dataset including data for bivalve mollusks and also with the dataset with the data for bivalve mollusks removed. The Asiatic clam (Corbicula fluminea) is in a different Family (Cobculidae) and Order (Veneroida), and is not a Unionid mussel. Corbicula has shown very similar sensitivity to ammonia compared to Unionid mussels however. Table B really represents a non-bivalve mollusk dataset, which would be appropriate where no bivalve mollusks of any kind are present. The FAV and CMC calculated without the data for bivalve mollusks is less than a factor of two different than the FAV using the entire dataset, which is 62.75 % higher. The 1999 ammonia criteria also provided separate recommended acute criteria based on observed sensitivity of salmonid species. The 1999 criteria

document also presented separate acute criteria that are approximately 50 % different between the salmonids present or absent acute criteria. The recommendations in the 2009 reassessment document providing separate criteria , calculated with and without a group of sensitive species is in keeping with precedents set in the earlier ammonia criteria for calculating separate criteria when one group of species appears to be especially sensitive to ammonia. This can be considered appropriate for ammonia criteria because of the relatively large database and the fact that ammonia toxicity is usually caused by mg/L concentrations as compared to µg/L concentrations as seen with many other toxic pollutants.

The implementation of criteria based on the dataset without freshwater mussels will require some way of determining whether or not a water body is not likely to contain freshwater mussels, before such criteria can be considered appropriate. The presence or absence of freshwater mussels in a water body may be difficult to ascertain without a survey that specifically targets this issue. Most monitoring techniques for collecting benthic macroinvertebrates such as the widely used EPA Rapid Bioassessment Protocol (RBP) are not designed or intended to regularly collect mussels that are often tightly imbedded in the sediment. Thus a typical RBP survey that did not collect freshwater mussels cannot be accepted as proof of absence of mussels in that waterbody. I would recommend a default of assuming the presence of freshwater mussels unless several surveys that specifically targeted the identification of mussels have been conducted on that water body. A brief review of biological survey data in Virginia showed that the presence of freshwater mussels was detected in 83 % of water bodies that had been surveyed with methods capable of finding mussels. However, in some waterbodies the presence of mussels was confirmed only after several surveys. Because the distribution of freshwater mussels is often dependent on fish for dispersal, it might be suspected that mussels may not be present in some headwater or intermittent streams where it can be demonstrated that fish are not present due to physical barriers preventing fish migration.

2.2.1.1. EPA Response

Since the publication of the draft 2009 ammonia criteria, additional toxicity testing has validated data on the effects of ammonia on sensitive freshwater gill-breathing snail species, which are relatively ubiquitous in distribution. Because these sensitive species (both mussels and snails) are widespread across the U.S., EPA is no longer recommending separate (i.e., bifurcated) criteria for aquatic systems with and without mussels. EPA is instead recommending one set of criteria that is reflective of the full freshwater toxicity database for ammonia, and protective of freshwater invertebrates overall, based on available data. The 2013 ammonia criteria document includes a section on site-specific criteria development, where appropriate (e.g., for sites where freshwater mussels are not present).

In addition, EPA has developed several supporting documents to aid states considering adoption of the 2013 recommended ammonia criteria. One of these documents, “Technical Support Document for Conducting and Reviewing Freshwater Mussel Studies for Development of Site-specific Water Quality Criteria for Ammonia”, is intended to help states determine whether sensitive freshwater mussels are present in their waters. Commonly-used mussel sampling methods will be described and an overview will be provided of various study approaches, considerations, and limitations, including real-life examples. EPA expects to publish this document on its website in 2013.

EPA is publishing two support documents on its website concurrent with the release of the final 2013 ammonia criteria document. “Flexibilities for States Applying EPA’s Ammonia Criteria Recommendations” provides a description of a number of implementation approaches available for state consideration, including the recalculation procedure for site-specific criteria derivation, variances, revisions to designated uses, dilution allowances, and compliance schedules. The document describes how each of these flexibilities fits within a state’s water quality standards adoption and implementation processes. The other document is a “Revised Deletion Process for the Site-Specific Recalculation Procedure for Aquatic Life Criteria”, which describes a recalculation procedure and includes a spreadsheet that may be used to derive site-specific water quality criteria for the protection of aquatic life in order to best reflect the organisms that reside at a specific site.

2.2.2. Peer Reviewer 2

Interestingly, we made a similar recommendation for this particular situation (i.e., the concept of with and without mussels-based criteria for ammonia) in that prior report (AWWQRP 2006, Chapter 4). I have attached a copy of the relevant chapter from that analysis, which may be useful, at least in reference to how this has been looked at by other parties.

In addition to a with and without mussel approach, the document could also include other “implementation” issues, such as evaluating additional site-specific information needed to provide in-depth technical recommendations. For example:

- Is there a definitive cause and effect relationship between ammonia sources and the diversity and abundance of mussels downstream of the source?
- How have the manageable sources (i.e., Agriculture, WWTP) of nitrogen changed over the last decade within the study waters in comparison to measured changes in mussel communities?
- Is there evidence of anthropogenic nitrogen loading that is equal to or greater than the natural occurrence of decomposition due to heterotrophic bacteria?
- What is the within-site spatial and temporal variability in NH₃ of the pore water and is this variability linked to season, storm events, temperature, or flow conditions.

In addition it would be advisable to include a statement to recognize the importance of developing “site-specific standards”, since mussels now play an important role and are the most sensitive species in the list. Site-specific standards reflect the reality of which species are needed to be protected in a particular waterbody, particularly with mussels that may be more abundant in some parts of the country, whereas in some other regions their presence can be very limited. In this case, some guidance on when and how to do mussel surveys to determine when to apply a “with mussels” or “without mussels” criterion would be a useful addition to the criteria document.

2.2.2.1. EPA Response

See response to peer reviewer 1 comment (Response 2.2.1.1 on p. 8). Additionally much of this comment regarding ammonia exposure and nutrient trend issues is beyond the scope of the ammonia criteria document.

2.2.3. Peer Reviewer 3

Attempting to differentiate freshwater ammonia criteria on the basis of mussels being present or absent is in my opinion a very tenuous and probably unsupported idea scientifically. First, the acute data indicate that *Corbicula* (Asiatic clam) is nearly as acutely sensitive as the most sensitive unionids. *Corbicula* is not a unionid and has a totally different life cycle than unionids. Furthermore, they are now widespread in the U.S. so it may prove fairly inefficient to separate criteria on the basis of presence of bivalves in general.

As an implementation issue, distinguishing between surface waters with mussels and those without is not nearly as straightforward as it is for salmonids or “cold water” species presence/absence because the latter are dependent on a fairly known, measurable water characteristic (water temperature regime). Also, while most states have cold water as well as warm water aquatic life uses in their water quality standards, I’m not aware of many (perhaps any) that have specifically a mussels use. The only example of which I’m aware is Virginia DEQ, which distinguishes different application of the chlorine standard based on whether federally listed mussels are known to be present in a given stream segment. They apply a “halogen ban” (no chlorine or other halogenated disinfection chemicals are allowed to be used for wastewater treatment) in those segments for those permittees that do not have intermittent dischargers and have discharges < 20,000 gallons per day. However, in this case the water quality standard was not

changed but rather a risk management approach was taken to reduce the major point sources of chlorine or other halogenated disinfection chemicals to the stream.

Mussels can and do occur (or have occurred) in a wide variety of habitat types and water quality regimes and so their presence is not often known without specifically sampling for them. Such sampling may need to be fairly intensive to “prove” the absence of mussels. In addition, mussels have been extirpated from a number of streams in the U.S. due to historic land use and habitat changes, toxic spills, and poorly treated wastewater and nonpoint runoff (e.g., see EPA 2002). If today, a state finds that mussels are not present but historically they were, what then should be the ammonia criteria? I believe such differentiations in criteria suggested in the RD be relegated to EPA’s recalculation of site-specific criteria (USEPA 1994). That procedure was designed to identify whether certain types of species are present or could occur in a given site or waterbody. This would be a more efficient and scientifically defensible way to handle the apparent sensitivity of freshwater mussels to ammonia.

2.2.3.1. EPA Response

See response 2.2.1.1.

2.2.4. Peer Reviewer 4

The document is silent on the appropriate methods for determining whether mussels are present or absent in a given receiving environment, and therefore which criteria should be applied. I expect that a guidance document will have to be developed to assist those applying the revised criteria. My expertise is in aquatic toxicology, while the question of presence or absence is based more in field-based aquatic ecology, particularly mussel ecology. As such I do not have detailed recommendations for the determination. I would, however, expect that some consideration would be given to a method that was able to deal with environments where mussels were once present, but have now been extirpated.

2.2.4.1. EPA Response

See response 2.2.1.1.

2.2.5. Peer Reviewer 5

With mussels presenting such an extreme sensitivity and complex life history in comparison to other groups, not only will their presence establish consideration for site variance demonstrations related to specific media, habitat or host fish interactions, but their spatial positions and ecological requirements against the backdrop of a stream’s ability to meet water quality standards- will require some unique approaches to:

- Short- term authorizations to discharge
- Compliance with permits that account for ammonia levels, but in streams or those segments that fail to meet water quality standards
- Designated mixing zones and how “shock effect” associated with ammonia will be considered
- The applicability of biocriteria and use of in-stream community assessments related to the concept of mussels present and absent that would incorporate differences between degree of colonization and distribution rates for fish, other benthic organisms of conventional use in stream surveys, and freshwater mussels. These considerations coupled with only a recent attempt to describe how certain stream characteristics may serve to structure mussel communities, will need to be carefully approached when and if such recommendation is made.

Since many of the ‘mussel present’ instances may involve listed species, the ability to qualify broad testing and culturing of suitable surrogate species for the time and seasonal limits related to any testing related to permit or enforcement actions, will likely present an additional set of laboratory qualifying or

certification considerations related to accessing suitably sited mussel species for testing. Laboratories listed in reference to the range of studies cited with this reassessment and that have implemented the new test standards (ASTM, 2006) have certainly advanced the methodologies needed for risk assessments and management decisions. However, this relatively limited number of research and agency laboratories will likely be challenged to fulfill the rigorous requirements of consistency with organism culturing and quality assurance that may be expected of more conventional test demands. While consideration of these outcomes may reach beyond the scope of the question for mussel presence, increased demand as a technical consideration would be expected to impact population availability.

Relative use of reference sites in determining relationships between areas of reduced freshwater mussel diversity and abundance and known significant ammonia sources will also likely become a technical consideration. Definitive cause-and-effect relationships have not been documented in such instances (Bartsch et al., 2003) and such considerations would require more detailed attention between the field exposure and effects data.

2.2.5.1. EPA Response

See response 2.2.1.1.

2.2.6. Peer Reviewer 6

I believe that EPA has adequately defended their position for when mussels were used or not used for the generation of the two phased acute criterion.

I do feel that EPA should include guidance that the lower mussel criterion should only be used when mussels are currently present in the water system and not historic or wished to be in the system. I can see States using the lower mussel numbers without a strong technical reason to use them because someone wished them to be in the system, etc. When States choose to go this way arbitrarily, it will weaken the acceptability of the criterion for everyone.

2.2.6.1. EPA Response

See response 2.2.1.1.

2.2.7. Peer Reviewer 7

Generally, water quality criteria are designed to protect populations. The first step in determining whether mussels are present or absent is to define a standard for the conditions describing that a population exists at a site. This is not a straight-forward exercise and can be quite difficult as seen in various site-specific criteria projects around the country. There must be guidance as to what parameters and measures of those parameters are to be used to define presence/absence of a population. Quantitative evidence of reproduction would be required at a minimum, as well as the presence of various developmental stages. The normal abiotic conditions documented for a species should also be compared to that of a particular surface water to determine whether a conclusion of presence or absence is supported. Certainly the availability of host species could influence the determination of absence/presence.

Once absence/presence is established EPA should only use mussel data to develop water quality criteria for waters where mussels are deemed present. EPA should ensure that conditions of the site where absence has been concluded have not been anthropogenically changed to a state where mussels cannot reproduce. Historic records of water quality and biological surveys would help support this determination. It is clear that EPA and the states will share the burden of the presence/absence determination and that there will be differences of opinion as to whether, when absence seems apparent, conditions changed due to human activity that could have been reasonably precluded or the species simply never successfully established a population at the site due to species-specific factors.

EPA should also determine whether simply the presence of mussels is important or the presence of specific species is important. Given that the sensitivity of mussel species in acute tests spans a six-fold concentration range the presence of only certain species could have a similar effect on the resulting water quality criteria; which then could have significant impacts on water quality management decisions. Further, different species have different life cycle durations. If individual species were used to develop criteria rather than the presence/absence of a taxonomic group the differences in life cycles may also impact criteria implementation.

2.2.7.1. EPA Response

See response 2.2.1.1.

2.3. Question 3: Is it scientifically defensible to exclude the glochidia data at this time due to the uncertainty of appropriate test duration time for this life stage? Do you believe there is an alternative approach to the use of this data that would be more scientifically sound?

2.3.1. Peer Reviewer 1

The position statement in Appendix B [A] of the reassessment document acknowledges several concerns about the use of toxicity data for glochidia and recommends that these data should not be used for criteria development at the present time. This position seems reasonable given the uncertainties regarding glochidia testing as outlined in Appendix B [A]. Glochidia, with such short-lived life stages and with variable lengths of natural viability in the water column, ranging from a few hours to several days, it is difficult to know how to use some of these data for developing criteria. For some species where the natural viability of glochidia, or where the actual exposure to the water column is short, the result of 48 hour tests may not provide a realistic exposure scenario, while a shorter exposure may be more important..

There are some additional short-term (6 and 24-hour) exposure data available that may provide some additional information on this issue. Table 2 of the Reassessment Document provides acute values based on 48-hour exposures for tests conducted with glochidia in several publications. For some of these same tests, additional information on 6 hour and 24-hour exposures are also contained in Wang et al. (2007b), in their table 3. These 6 and 24-hour data indicate that glochidia are less sensitive to ammonia during 6 and 24-hour exposures than during a 48 hour exposure. For the species where glochidia may only be viable or exposed to the water column for a few hours, the results of 6 hour and 24 hours tests may provide useful information and these data could be normalized to the proper standard conditions and included in Table 2, to provide additional information for comparison. As reported in Wang et al. (2007b); the data suggest that at 6 to 24 hour test durations, the sensitivity of glochidia for most species of freshwater mussels is less than or equal to the sensitivity of the juveniles of the same species at 96 hours. This provides some support that the revised ammonia acute criteria (based on 96 hour data for juveniles) may also provide protection to glochidia at exposure durations of 24 hours or less. The fact the acute criterion is applied as a one-hour average should also provide some additional level of protection.

2.3.1.1. EPA Response

EPA published the 2009 draft revised criteria for ammonia taking into account the most current scientific information at that time, including glochidia toxicity test data and recommended test methods (ASTM, 2006). At that time, EPA (consistent with external peer reviewers' comments at that time) concluded that information was insufficient to determine whether the glochidia toxicity tests were scientifically sound and appropriate for quantitative use in criteria development. Specifically, the appropriate duration of the tests (24, 48, or 96 hours) was uncertain because it was unclear how duration related to the viability of the short parasitic life stage of glochidia and its ability to successfully infect a fish host. Since that time,

a study by Bringolf et al. (2013) has shown that a maximum test duration of 24 hours is appropriate for glochidia, corresponding with the ecologically-relevant endpoint of infectivity of a fish host. As a result of this new information, EPA has included glochidia toxicity test data in the 2013 ammonia criteria dataset, provided that the maximum test duration was 24 hours and control survival of glochidia at the end of 24 hours was at least 90 percent. In addition, to account for species of mussels whose glochidia might not be expected to be viable at 24 hours (i.e., potentially mantle lure strategists), EPA examined available tests with glochidia that were conducted for 24 hours that included testing for viability at 6, 12, and 18 hours. If the viability was less than 90% at 24 hours in the control animals, then the next longest duration less than 24 hours that had at least 90% survival in the control, was also considered acceptable for use in deriving the ammonia criteria. With respect to ammonia toxicity, the glochidia were not consistently more sensitive than juvenile mussels. For example, for the oyster mussel the glochidia were more sensitive than the juveniles by more than a factor of 2, but for fatmucket mussel the juveniles were more sensitive than the glochidia by more than a factor of 2, and for the other 4 mussel species with acute data for both glochidia and juveniles the acute values for both life stages were different by less than a factor of 2. Since the glochidia were not always the more sensitive life stage and/or the difference between the glochidia and juvenile toxicity values were less than a factor of 2 per the 1985 Guidelines, data for both glochidia and juvenile mussels was combined in the derivation of the CMC. Thus, glochidia and juvenile mussel data were included and averaged in the criteria derivation to protect these sensitive life stages.

2.3.2. Peer Reviewer 2

Given the rationale provided, in combination with EPA criteria development guidance, general ecotoxicological principles, and my own views on general invertebrate biology, I agree it is scientifically defensible to exclude glochidia data in ammonia criteria development for the following reasons.

- First, the document notes that despite having four acute toxicity studies that consider five different mussel species (with the glochidia stage), with the majority of test having a duration of 2 days, there are still several questions to be answered when testing these particular life stages:
 - How long do they stay attached to the host?
 - How long do they stay in conglomerates and does this jelly (for those species) protect them from exposure to contaminants?
 - If they sink to sediment how long do they stay there in the conglomerate before transforming to the juvenile mussel?
 - I agree with the authors that these valid questions and also note that the questions will probably have different answers for different species of mussels, which makes the data of even more uncertain general ecological relevant and, therefore, use of glochidia data even more problematic and difficult to apply generally to this group for criteria development purposes.
- Second, concerning the mechanisms of ammonia toxicity in this particular stage of development there is also a long list of unanswered questions: would glochidia be able to take up ammonia if available in the host?, what is the mechanism of glochidia exposure to ammonia when they are released in conglomerates? - Again – great uncertainty with unknown ecotoxicological significance.
- Third, based on the results from those four studies it seems that if the glochidia acute data were included in the calculations, they would dramatically reduce the acute criteria and probably be over conservative for other fresh water species without actually knowing if the information has any ecological relevance.

Given these uncertainties – which are well articulated in the draft document, there appear to be two scientifically defensible approaches: 1) Do not use the glochidia acute toxicity data 2) Do a final acute criteria calculation “with and without glochidia”. The first approach would be the easiest to avoid confusion and possible errors in development of safe levels for ammonia. The second approach would give flexibility – although it would be necessary to include a strong caveat to a “with glochidia” value, given the questions noted in the criteria document and my thoughts above.

2.3.2.1. EPA Response

See response 2.3.1.1.

2.3.3. Peer Reviewer 3

I agree with the RD that glochidia data should not be included in criteria derivation at this time but not entirely for the reasons given by EPA. First, a female mussel often produces hundreds or even thousands of glochidia in a single season. This helps counteract the fact that a high percentage of the glochidia will not survive because they have not encysted on a host fish within a reasonable amount of time. As many researchers have documented, glochidia are incapable of surviving and maturing on their own—they require nourishment from a host organism typically within days of being produced. It could be argued that basing criteria on a 50% reduction in glochidia survival would be similar to basing criteria on a 50% reduction in egg survival of other species that broadcast an abundance of eggs (e.g., marine barnacles) in which only a small fraction of the larvae are likely to survive and become established under natural, unstressed conditions. I am not aware that EPA has ever based acute criteria on egg survival (even fertilized egg survival) and I believe this might contradict EPA’s criteria development guidelines for establishing acute criteria.

A second reason I don’t think glochidia should be included at this time is that the majority of mussel species at risk today (i.e., threatened, endangered, or species of concern) appear to have relatively narrow specificity in terms of a vertebrate host and have evolved structures (e.g., lures) or behaviors that enable very rapid encysting in the host (e.g., see Rogers-Lowery and Dimock 2006). Of the mussel species actually used in glochidia testing that are discussed in the RD, most of them have evolved adaptations to enable rapid infestation of a host (often < 6 hours), resulting in a very short free-living glochidia stage. Thus, for the majority of mussel species at risk, exposure periods > 24h are probably inappropriate, regardless of whether control survival is > 90%.

A third reason is that while the ASTM glochidia test procedure appears to be reasonably robust in terms of intra and inter-laboratory testing, the procedure is still a “Guide”, not a “Method”, which indicates that there are still many aspects requiring expert judgment, trial and error, and research, especially in terms of culturing and maintaining glochidia under laboratory conditions. Also the inter-lab study reported by Wang et al. (2007) used the same water source and glochidia stocks for all labs. This would tend to underestimate the true interlab variability, where each lab uses their own source of glochidia (and techniques for obtaining glochidia) as well as water source. In addition, at the present time there are still relatively few laboratories that have performed this test (e.g., compare with daphnia or amphipod tests). Much of the glochidia data under consideration were produced by a few laboratories. This may be acceptable for tests that use species closely resembling those for which a standard EPA or ASTM method has been developed and used extensively; e.g., an acute test with a different species of minnow or Daphnia but the identical test design and organism life stage as used for the standard fathead minnow or Daphnia acute test. However, such is not the case with the ASTM glochidia test, which uses a fairly different test design and the endpoint is dependent on sensitivity to a sudden stress (i.e., rapid valve closure to NaCl exposure). While the ASTM procedure appears to address these concerns for the most part, there is still some uncertainty in my mind regarding defensibility of the test endpoint. As noted in the ASTM method, exposures beyond 24h may not achieve 90% control survival, indicating the precarious nature of keeping this life stage alive without a host. Given this, is the apparent increase in

ammonia sensitivity between 6 and 48h a true difference in sensitivity or is it an artifact of the method (i.e., glochidia are under increased stress). I also question the ecological importance of having valve closure within one minute or less to the introduction of a NaCl solution. I'm not altogether sure this endpoint is similar to other acute endpoints relied upon by EPA in their water quality criteria. In addition, the ASTM Guide points out that live or dead mussels could be open or closed at a given time and some may respond more slowly to the NaCl shock than others (Kernaghan et al. 2005). If there was a more direct way to determine mortality of glochidia to ammonia exposure, for example, that would be preferable in terms of using such data for criteria development.

I wonder if better use of glochidia survival data couldn't be made by considering the level of juvenile recruitment necessary to maintain a viable mussel population, similar to the EPA fish larval recruitment model for dissolved oxygen in coastal waters in the Virginian Province (USEPA 2000). Perhaps a similar type of framework could be used based on the mussel species (or, if feasible and defensible, genus or even family), when glochidia are typically produced in a given region, and a function specifying ammonia toxicity to glochidia given ambient temperatures and pH at that time. The modeling framework could be made general so that a user could input necessary temperature and pH data, recruitment timing (e.g., season, month) and certain assumptions regarding host abundance. This would not be unlike the larval recruitment model EPA has developed for dissolved oxygen with the exception that the abundance of another species (i.e., a host) is not needed or relevant for the dissolved oxygen model. However, if such a model was considered (albeit simplistic but using conservative assumptions based on expert malacologist judgment), it could put glochidia toxicity testing results into an ecological context in a more useful way.

2.3.3.1. EPA Response

See response 2.3.1.1.

2.3.4. Peer Reviewer 4

Exclusion of the glochidia data, as detailed in Appendix A, is fully appropriate and well explained. Until tests are completed to determine (on a species by species basis) the duration of the period of longevity (and by inference viability) for glochidia, LC50s obtained with those life stages will be suspect. I am unaware of any other approach which would be more scientifically valid

2.3.4.1. EPA Response

See response 2.3.1.1.

2.3.5. Peer Reviewer 5

Information from tests with glochidia can be supportive of the consideration for site variance information, but should not be used to provide stand alone estimates in the absence of supportive acute juvenile estimates. Multiple lines of evidence comprised of testing with more than one mussel life stage that include comparisons with glochidia responses, would seem more scientifically sound, and may remove much of the seasonal dependence upon culturing and provision of test organisms with less invasive techniques in instances mentioned in the above mentioned technical considerations.

2.3.5.1. EPA Response

See response 2.3.1.1.

2.3.6. Peer Reviewer 6

Not only is it scientifically defensible, I believe that it is the right way to go. Too much is unknown with regards to the biology of glochidia and the associated toxicology or appropriate methods to evaluate the toxicity of various chemicals for this life stage.

2.3.6.1. EPA Response

See response 2.3.1.1.

2.3.7. Peer Reviewer 7

The alternative question would be: Is it scientifically defensible to include glochidia data at this time due to the uncertainty of appropriate test duration time for this life stage? Again, EPA has not defined standards in the Guidelines for this determination other than requiring acute tests be at least 48 hours in duration and test organisms not be fed (except for mysids). These requirements were not developed to address the duration of the life stage unique to glochidia. Toxicity data for glochidia appear to show that duration of tests with this life stage significantly affect the test results when tests are 24 hours in duration or longer. Therefore it seems that the glochidia are stressed in some fashion when tests meet the 48 hour requirement of the Guidelines. The test design for water quality criteria development is intended to measure stress due to the treatment (ammonia) independent of all other stressors. Tests of 48 hour duration and more fail to meet this goal. Since tests of duration less than 48 hours do not meet the Guidelines' requirements but tests of duration equal to or greater than 48 hours provide biased results it seems that a new standard for test duration unique to these mussels must be first developed. However the literature indicates a great deal of diversity among freshwater mussels in terms of their glochidia ecology. The extent of this issue demands either more detail in this criteria document or more opportunities for site-specific ammonia criteria implemented by the States for only the mussel species present. The States will also require guidance on how to develop site-specific ammonia criteria based on the issues unique to glochidia. To support the scientific defensibility of EPA's decision it would be appropriate to first develop guidance for the appropriate test duration to be used for developing water quality criteria relative to the duration of the life stage. Currently acute test duration commonly is within an order of magnitude of life stage duration. To accomplish this goal for glochidia a test duration of 12 hours or less may be necessary. Such a guideline would help defend the scientific defensibility of using glochidia data for developing water quality criteria.

One should conclude that EPA's decision to not use glochidia data is logical and correct without more information. There is a great deal of uncertainty associated with test results using glochidia. The question that must be answered is whether the level of uncertainty realized by using glochidia data is so great that it compromises the integrity of the resulting water quality criteria. This, in fact, seems the case; however it should be stressed that scientific defensibility should not be based on judgment but minimum standards of quality. Without these standards of quality the magnitude of uncertainty is unknown and probability of erroneous actions is elevated.

2.3.7.1. EPA Response

See response 2.3.1.1.

2.4. Question 4: Regarding the proposed approach to glochidia data in the 2009 draft position statement as it relates to ecological relevance and practicality - Is the approach a scientifically defensible principle for structuring the population exposure duration problem and designing further research to quantify such a duration?

2.4.1. Peer Reviewer 1

I believe that the approach outlined in appendix B [A] is defensible. Given the significant uncertainties concerning the appropriate exposure durations for glochidia as discussed in the appendix B, attempting to use toxicity data currently available for glochidia to derive defensible criteria at the present time appears to be untenable. This would be especially important if the final criteria were significantly influenced by

glochidia data with the recognized uncertainties about practical ecological importance. Water quality criteria should be based on the best information available in order to gain acceptance and avoid continued uncertainty. Additional information as discussed in appendix B would be needed to provide sufficient confidence in utilizing glochidia data to derive water quality criteria.

2.4.1.1. EPA Response

See response 2.3.1.1.

Also, Bringolf et al. (2013), via a Final Completion Report submitted to EPA Region V, completed a study to determine the appropriate duration of toxicity tests with glochidia of native freshwater mussels, which has been submitted to Environmental Toxicology and Chemistry for peer-review under the same title and with the following list of authors: Andrea Fritts, Christopher Barnhart, Megan Bradley, Na Liu, W. Gregory Cope, and Robert Bringolf. As stated in the Executive Summary to this report, the specific objectives of the study were to determine:

“ if the duration of glochidia viability, defined as the ability to close in response to salt (NaCl) solution, is equivalent to the duration of infectivity [defined as]- the ability to attach to a host fish and metamorphose successfully into the juvenile stage;

if the duration of infectivity of glochidia deposited in natural stream sediment is similar to the duration observed in water-only exposures; and

the probability that glochidia deposited onto the sediment surface will infect host fish, using representative mussel-host species pairs.”

Importantly, the results indicate that the glochidia viability endpoint [of $\geq 90\%$] as recommended in the current ASTM guidance (E 2455-06) does adequately represent metamorphosis success of glochidia, but with a maximum test duration of 24 hours. Thus, glochidia viability is an ecologically relevant, and scientifically-defensible, endpoint adequate for use in standard toxicity testing protocols for the parasitic (larval) life stage of freshwater unionid mussels, especially for species with a defined duration of viability of ≥ 24 hours. For those species with little or no such data regarding duration of viability, viability can also be assessed at an intermediate time point such as 6, 12 or 18 hours, and if viability is $\geq 90\%$, these data are also suitable for criteria development.

2.4.2. Peer Reviewer 2

The proposed approach will give researchers and scientific community time to come up with some answers and provide relevant information concerning the appropriate time of duration for glochidia exposure testing – if such testing is determined to be necessary for this short-lived, parasitic life stage. However, if this approach is proposed, it will also create a limbo period, in which this issue will have to be addressed while decisions are made for application of a criteria that some may fight as “non-protective” since the criteria document has “acknowledged” that a sensitive life stage was not considered!

I’ll be the first to admit it is a difficult to task to structure a criterion that includes population exposure approach to criteria development when dealing with a wide variety of fresh water mussel species that have different life spans (few months to several years), not to mention a glochidia life-stage that can be released individually or in conglomerates, and that can get attached to fish, rocks, sediment or plants.

Such variability could mean that for the particular issue of including glochidia data or not, we perhaps should be treating these mussels individually, by genus or species, rather than thinking of them simply as the “mussel” group. Regardless, it will take a considerable amount of time and effort to develop some way to make “uniform” relevant exposure toxicity testing time for the glochidia life stage that will be relevant for all if not the majority of fresh water mussels.

2.4.2.1. EPA Response

See response 2.3.1.1.

2.4.3. Peer Reviewer 3

I am not sure I concur entirely with the proposed approach for further research discussed in Appendix A of the RD. I do agree that it would be useful to have better information regarding the natural life expectancy of free-living glochidia in nature for several representative species covering the range of different reproduction strategies observed. Such information would help indicate whether a 6, 24, or 48h exposure is warranted. I do not, however, understand EPA's proposal specifying the free-living duration based on 95% of the glochidia that attach to a host. This appears to be a fairly restrictive proposition and it is not clear to me how this would be determined. As Rogers-Lowery and Dimock (2006) observed, the encapsulation process in fish varies with the species and organism history of exposure to mussel infestations. Many researchers have demonstrated poor encapsulation rates of glochidia of certain mussel species (including several listed as threatened or endangered) with several common fish species (e.g., Dodd et al. 2006). In other words, the number of glochidia that encyst on a host has as much to do with the host as it does glochidia viability in and of itself. Somehow, there would need to be the presence of a known host with the glochidia to determine an answer to EPA's proposal. However, viable host species are not known for many mussel species. Therefore, this evaluation would require testing with a few mussel species for which host species are known with certainty.

2.4.3.1. EPA Response

See response 2.3.1.1.

2.4.4. Peer Reviewer 4

I am not a mussel ecologist, but the approach taken, and its implications for further research appear valid. Having said that, it is likely that modifications to the approach will have to be made as such research proceeds. It is also apparent that there is sufficient species variability in the life habits of glochidia that a single approach will not fit all species.

2.4.4.1. EPA Response

See response 2.3.1.1.

2.4.5. Peer Reviewer 5

The support of life history information and exposure information can support this contention, but the pertinence of this endpoint would continue to be challenged for its scientific defensibility as relates to what we currently understand of exposure durations related to this life stage and variety of adaptations.

2.4.5.1. EPA Response

See response 2.3.1.1.

2.4.6. Peer Reviewer 6

See comment for Question 3, the same comment is relevant for Question 4.

2.4.6.1. EPA Response

Noted, thank you.

2.4.7. Peer Reviewer 7

As discussed earlier the uncertainties of using juvenile mussel data for deriving water quality criteria are significant, but EPA has identified additional concerns associated with tests of glochidia. The sensitivity of glochidia as a function of exposure duration, within the constraints of the Guidelines, precludes the use of glochidia data to derive water quality criteria. This is primarily due to the relatively short duration of

instream exposure that can occur with this life stage and the sensitivity of glochidia to that duration of exposure. The issue of exposure duration instream versus that of lab toxicity tests is critical to the use of all toxicity data in deriving water quality criteria and is not unique to freshwater mussels and their glochidia life stage. Decapod crustacean larvae are sensitive to duration of test exposure without food for durations required by the Guidelines. The literature refers to the “Point of No Return” or the PNR (Klaus Anger, various papers), where larvae will not survive beyond a certain time period without food even if food is provided to the larvae. Toxicity tests of duration approaching the PNR are likely stressing the test organisms in addition to the toxicant in question. The same can be said of glochidia tests with durations of 48 hours and more and likely durations even greater than 24 hours.

Development of criteria must consider all test factors that might bias test results in comparison to responses that are expected instream, and exposure constitutes one of the most important factors. Exposure is a function of frequency, magnitude and duration. The default frequency assumption for water quality criteria toxicity tests is that exposure is continuous. Magnitude of exposure is usually dependant on concentration, which may also be a function of other factors like pH, alkalinity, DOC, etc. The magnitude of exposure for glochidia will also be affected by the species-specific conditions unique to glochidia of different species. For example, some glochidia are released into the water column singly; some are released in groups encased in a gelatinous shell where exposure would be mitigated. Duration of test exposure should be related to the duration of the life stage, unless it can be demonstrated that a different test duration is representative of that life stage. This is a fundamental of the Guidelines although it is not clearly delineated. In the case of glochidia 48 hour or longer test durations have been shown to result in responses not expected instream. This is a defensible reason to reconsider use of the acute glochidia data to derive water quality criteria and defines well the research needs for this taxonomic group and life stage. Shorter duration tests will be necessary to provide reliable test results for deriving water quality criteria.

2.4.7.1. EPA Response

See response 2.3.1.1.

3. HYALELLA AZTECA POSITION STATEMENT AND PROPOSED RATIONALE: (see Appendix B)

The EPA workgroup developed a position statement and proposed supporting rationale describing the concerns over using *Hyalella azteca* toxicity test data in criteria development due to the uncertain health of the test organisms in different test water composition at that time. The rationale defined the specific concerns and uncertainties supporting the recommended exclusion of the *Hyalella* data from use in criteria derivation, at that time.

3.1. Question 5: Are the position statement and supporting rationale regarding use of toxicity data for *Hyalella azteca* in criteria development reasonable and scientifically sound recommendations?

3.1.1. Peer Reviewer 1

Yes. The EPA Guidelines for Deriving Numerical National Water Quality Criteria for the Protection of Aquatic Organisms and Their Uses, in section IV. 5. H specifies that in order to be useful for calculating a Final Acute Value, there must be a relatively good agreement of the data within a species. Generally “if the acute values differ by more than a factor of 10, some or all of the values probably should not be used in the calculation” of a Species Mean Acute Value (SMAV). This applies to the available data set for *Hyalella azteca*, where the normalized acute values listed in Table 2 range from 1.58 mg/L to 83.9 mg/L,

differing by as much as a factor of 53. This wide range of acute values for *H. azteca* required more careful review before determining what if any of the data for this species is acceptable for criteria development. EPA has conducted an additional review of these data and has determined that there are significant uncertainties with data for this species, enough to preclude their use in deriving criteria. The position statement provides a summary of the information known concerning possible factors that could cause the acute toxicity values for *H. azteca* to be affected by some water quality characteristics not previously recognized or considered and not fully understood at this time. Based on the current information available, it appears reasonable that until these factors are better understood, the data for *H. azteca* should not be used in the calculation of a SMAV or a FAV.

3.1.1.1. EPA Response

*In the 2013 final ammonia criteria document, EPA noted that it decided to use certain data for Hyaella azteca based on recent scientific studies regarding the water chemistry constituents required to maintain the health of these test organisms for their use in aquatic toxicity testing (described in detail below excerpted from the 2013 ammonia criteria document, p. 57). The position regarding use of Hyaella azteca data in the draft 2009 document was specific to the ammonia toxicity data only; the acceptability of Hyaella data for use in aquatic life criteria for other pollutants will be evaluated on a study by study basis at this time. It is worth noting that in the 2013 criteria document chronic dataset, Hyaella was found to be less sensitive than 12 of the 16 genera tested. Thus, the *H. azteca* data was not a driver in the chronic criterion magnitude calculation.*

28-day toxicity data for Hyaella azteca

Literature data indicate that the response of Hyaella azteca is influenced not only by pH, but also by sodium concentration in the dilution water. Borgmann and Borgmann (1997) demonstrated that increasing sodium decreased the toxicity of ammonia to Hyaella, and applied these findings to explain differences in toxicity observed by Ankley et al. (1995), which were originally attributed to water hardness. Further unpublished experiments by EPA's Office of Research and Development confirm Borgmann's assertion that sodium concentration plays a key role in determining the acute response of Hyaella to ammonia (personal communication, D.R. Mount, EPA, ORD). Because sodium is not known to affect ammonia toxicity to other species, this criterion does not consider sodium concentration, and this variation is not explicitly addressed. For purposes of deriving a GMAV for Hyaella in the 2013 ammonia criteria document, tests were selected that had a moderate sodium concentration (e.g., "moderately hard" water tests from Ankley et al. 1995, see Appendix A), and tests with extremely low sodium concentrations were excluded (e.g., "soft" water tests from Ankley et al. 1995; data from Whiteman et al. 1996, see Appendix J). The available acute data for ammonia did not include tests conducted in natural waters with a sodium concentration below about 3 mg/L; at that sodium concentration, the acute values for Hyaella were near the FAV reported in this document. Whether acute toxicity of ammonia to Hyaella would occur below the FAV in waters with less than 3 mg/L sodium is unknown (Appendix H).

*For the 2013 chronic criterion, EPA re-evaluated the available data for Hyaella azteca based on recent research regarding the appropriate test conditions, including water chemistry (e.g., appropriate concentrations for specific ions such as chloride) and feeding regimes. The concentrations of sodium are important to *H. azteca* health as discussed previously and the sodium concentrations in the chronic test used in the CCC represent approximately the mid-range of U.S. waters. Based on this re-evaluation, EPA determined that certain tests met the new recommended conditions that would support healthy test organisms, and accepted those data for use in the calculation of the CCC. The specific tests used were from Borgmann (1994); details on these tests are included in Appendix H under Chronic Toxicity Tests with Juvenile Hyaella azteca. As a result of inclusion of acceptable *H. azteca* data, the minimum data requirement (MDR) for a benthic crustacean is fulfilled for the chronic criterion provided in this document. The GMCV of 29.17 mg TAN/L ranks Hyaella azteca as the thirteenth (of 16) most sensitive GMCV.*

3.1.2. Peer Reviewer 2

I support the position statement based on the workgroup's review of a number of toxicity tests on *Hyaella*, referenced in the rationale. I can't help but wonder if this recommendation holds for *Hyaella* test data for other EPA criteria documents, as well – for example, *Hyaella* is the most sensitive species in EPA's chronic cadmium database (EPA 2001). Is this position statement generic to all EPA criteria databases, or simply this ammonia update?

I should note that our prior review of the study also recommended removal of *Hyaella* from the ammonia database, primarily because of the poor control performance (AWWQRP, 2006), which it now appears may be due to the ionic balance issue noted by the position statement. So, I support non-use of *Hyaella* toxicity data for this ammonia update.

Most of the recommendations provided have scientific basis. It appears that ongoing studies will provide quality data that will widen the knowledge for water quality requirements for *H. azteca* husbandry. However, the statement concerning the complications due to possible *H. azteca* genetic or taxonomic diversity is not relevant because the way this information is used in criteria development, which only considers appropriate data at genus level (i.e., GMAVs) not new species or subspecies. In addition, the issue of a wide range of surface waters inhabited by *H. azteca* could be easily addressed setting a site-specific standard, which will consider only the species present at a particular site. Thus, inclusion of this information may add confusion rather than help explain a decision on non-use for criteria development.

3.1.2.1. EPA Response

See response 3.1.1.1.

3.1.3. Peer Reviewer 3

I have reviewed the *Hyaella* data in a previous outside peer review for EPA and agreed with EPA's concerns regarding *Hyaella* test data for ammonia. I agree with the position statement in Appendix B that *Hyaella* toxicity data should not be used in water quality criteria development. There is a clear need to resolve water quality requirements and other methodological issues for this species in water-only tests.

3.1.3.1. EPA Response

See response 3.1.1.1.

3.1.4. Peer Reviewer 4

The exclusion of *Hyaella azteca* data, as summarized on page 13, is fully defensible. I would suggest that while all of the points are relevant, the pivotal consideration is the impact of chloride and bromide ion on the viability of the species and the potential interaction of the concentration of these ions with ammonia toxicity. The *Hyaella* data will be suspect until that information is obtained.

3.1.4.1. EPA Response

See response 3.1.1.1.

3.1.5. Peer Reviewer 5

Those points are very critical given the consideration for osmoregulation by this organism and the points cited in the workgroup's review. This seems well illustrated in the background material furnished for this review.

3.1.5.1. EPA Response

See response 3.1.1.1.

3.1.6. Peer Reviewer 6

I believe that the position statement and supporting rationale is appropriate and justified.

3.1.6.1. EPA Response

See response 3.1.1.1.

3.1.7. Peer Reviewer 7

The position statement is written to address the use of *H. azteca* data for the ammonia update; however the issues presented would likely be important for many other water quality criteria using data for this species. Therefore the position statement is not reasonable and scientifically sound if *H. azteca* data are considered within the context of all water quality criteria calculations. There is a concern that decisions such as this do not have associated rules as normally defined in the Guidelines. EPA is being reasonable and responsible in considering factors that may bias the development of water quality criteria in this case, but it is unclear whether EPA has a process to identify such factors as new information is collected over time. Perhaps there was insufficient data to identify the *H. azteca* issues when data for this species was considered for use, but EPA needs to develop and implement a system to periodically review data for issues such as this and publish the system as part of the Guidelines.

Based on this comment the position statement may be reasonable but the scientific basis of the recommendation is questionable without a well defined system of review. Such a system should include review of data relative to standards of quantity and quality required to make such an assessment. EPA needs to determine how much information is necessary to make this type of recommendation as well as the qualitative nature of the data (measured vs. nominal concentrations, static vs flow through tests, etc.) Without a system it will become increasingly difficult to be consistent in making these types of decisions as more variances arise.

3.1.7.1. EPA Response

See response 3.1.1.1.

4. CHRONIC FRESHWATER CRITERION:

4.1. Question 6: Are the toxicity tests and other studies used to derive the criterion scientifically defensible for such use? Are you aware of other relevant data that were not used?

4.1.1. Peer Reviewer 1

Yes. The toxicity studies referenced all appear to be good quality studies and provide useful information on the chronic toxicity of ammonia to aquatic organisms. The use of EC20 values calculated by regression analysis is excellent and provides a consistent estimation of low-level effects under chronic conditions.

All toxicity data relevant to ammonia toxicity that I am aware of are included in the references and were considered in the reassessment document.

4.1.1.1. EPA Response

Thank you for your comment.

4.1.2. Peer Reviewer 2

While, the toxicity tests used to derive the criteria are generally scientifically defensible (see discussion below); I am uncomfortable with the calculation because there is insufficient data to evaluate intra-species comparisons using the guidelines established for deriving water quality criteria. More specifically, the 2009 data set still does not fully meet the EPA guidelines for developing water quality criteria. A salmonid representative has been added to the data set (which was missing in the 1999 document), but a representative from the Class Insecta is still missing. It appears that the option for using acute-to-chronic ratios for determining chronic criteria would be more relevant than simply using a “hypothetical GMAV for insects”.

One of the studies noted in Question 1 may provide additional relevant chronic data.

It is also important to point out that there are only two studies (Anderson et al. 1978, Sparks and Sandusky 1981) available for *Musculium* genus. Interestingly, those two studies provide substantially different estimates of what level of ammonia is protective to fingernail clams – creating uncertainty in the true chronic value. A closer analysis of these two papers reveals that while the later Sparks and Sandusky study was designed to confirm Anderson et al’s findings, the researchers faced some serious methodology deficiencies, as the authors recognize: “Although it would have been desirable to perfect the culture methods first and then employ them in the toxicity tests, methods were developed as testing proceed because of limited time” (Sparks and Sandusky 1981). As part of their study design, Sparks and Sandusky used clinoptilolite to remove ammonia, but note this compound could also have remove potassium and chlorinated hydrocarbons – which raises the question that toxic effects observed in *Musculium transversum* in this particular study may be also attributed to other confounding factors in the test water and not exclusively to ammonia exposure. In addition, authors conclude that also biological processes (i.e., naturally occurring bacteria that converts ammonia into a relatively non-toxic nitrate) may be involved in ammonia removal rather than physico-chemical processes. Thus, it would be advisable to re-analyze the relevance of these studies for this update. As it stands, I cannot fully support calculating a GMAV for this fingernail clam using these two studies.

Another issue that is relevant to mention is the continued inclusion of a temperature relationship for chronic ammonia criteria originally used by EPA (1999), which was derived from a single study. It appears the EPA used the Arthur et al. (1987) study, which evaluated acute ammonia toxicity to 14 species (9 invertebrate and 5 fish species), to incorporate temperature dependence into the chronic equations. There are three problems with incorporating temperature in the chronic relationship, but not acute. First, there is lack of chronic ammonia toxicity studies for fish or invertebrate species to demonstrate the relationship between temperature and chronic ammonia toxicity. Second, the rationale for deriving an invertebrate chronic temperature slope from acute data (i.e., fish and invertebrates) is unclear. Finally, the assumption that chronically exposed invertebrates will have similar temperature dependence compared to acutely exposed individuals is based on the Thurston et al. (1984) study and those authors indicated that a 96-h period is insufficient to determine an acute toxic concentration of ammonia for insects. Consequently, it appears there are no appropriate data to support incorporation of a temperature component to the chronic ammonia standards, just as EPA concluded with the acute equations. The position statement notes that the temperature-dependent toxicity model now only applies to invertebrates.

I have to admit I was hoping the actual equations, inflection points, and pH and temperature relationships would be thoroughly re-evaluated as part of this update, given the abundance of new data included not in the 1999 document. I would strongly recommend re-evaluation of the chronic temperature component of the equations rather than just accept the 1999 update evaluation. Although, I have to admit, the draft document doesn’t actually provide any equations to test – just recalculations of acute and chronic endpoints. This is another disappointment.

4.1.2.1. EPA Response

All 8 minimum data requirements per the 1985 Guidelines (see Sections III.B and VI.A) are now considered fulfilled in the 2013 ammonia chronic dataset. Following a re-analysis of the salmonid and insect data, both taxa are represented in the 2013 ammonia chronic criterion dataset (summarized on p. 56-61 of 2013 ammonia criteria document, respectively; also see App. B chronic toxicity data table, p. 131).

For the requirement of at least one freshwater animal species in the family Salmonidae, the 2013 criteria document re-considered the results reported in six chronic ammonia toxicity studies with *Oncorhynchus mykiss* and *Oncorhynchus nerka* that had been previously evaluated in the 1999 criteria update, as well as two additional “new” studies involving *O. mykiss* (as reported in Brinkman et al. 2009) and *O. clarkii* (as reported in Koch et al. 1980). Decision criteria were established (see p. 60 in the 2013 ammonia criteria document) that allowed EPA to include all available and reliable chronic toxicity test data that should be used to derive a Genus Mean Chronic Value (GMCV) for this recreationally, commercially and ecologically important taxon. The new GMCV calculated for *Oncorhynchus* in the 2013 criteria document is 12.02 mg TAN/L, which is the geometric mean of the three SMCVs derived separately for *Oncorhynchus clarkii* (25.83 mg TAN/L), *O. mykiss* (6.663 mg TAN/L), and *O. nerka* (<10.09 mg TAN/L) (see Appendix B in the 2013 criteria document). Based on this value, EPA identified *Oncorhynchus* as being the seventh most sensitive GMCV in the 2013 chronic criteria dataset (see Table 4 in the 2013 criteria document).

For the requirement of an insect in the chronic criteria dataset, the 2013 criteria document re-considered the results reported for the stonefly, *Pteronarcella badia*, from Thurston et al. (1984a) which had also been previously evaluated in the 1999 criteria update. Upon further consideration of those data, EC20s for 30-day nymph mortality were calculated for field collected *P. badia* for two separate partial life cycle tests in consecutive years, in order to develop a GMCV for that insect species. The normalized EC20 for the test conducted with *P. badia* collected from the Gallatin R. was 207.0 mg TAN/L, and was 26.27 mg TAN/L for the test conducted with *P. badia* collected from the Rocky River (Appendix B in the 2013 criteria document). The geometric mean for these two tests is 73.74 mg TAN/L, which lies well above (10 times) the GMCVs of the four most sensitive species in the 2013 chronic criteria dataset.

In addition, chronic data for *Hyaella azteca* are included in the 2013 ammonia criteria as discussed above in response 3.1.1.1. Moreover, an ACR is used with acceptable acute toxicity data to populate the MDR for an insect or other phylum not represented (i.e., Annelida).

With respect to re-evaluating the pH and temperature ammonia toxicity relationships established in 1999, EPA conducted a search in August 2010 of its ECOTOX database for additional ammonia and ammonium studies published between 1985-2009 which were previously not considered or used for developing the pH and temperature slope equations included in the 1999 AWQC document. Four acute studies were found that could be used to evaluate the pH-TAN relationship (Straus et al. 1991, Hickey and Vickers 1994, Borgmann and Borgmann 1997, and Wang et al. 2008), and two acute studies were found that could be used to evaluate the temperature-TAN relationship (Hickey and Vickers 1994, Sarkar 1997). After review, none of these studies were used to calculate the pH or temperature equations for TAN toxicity provided in the 1999 AWQC document for various reasons (e.g.; test duration, non North American species, TAN added as a formulation, publication subsequent to equation development). No additional chronic studies could be found to evaluate the chronic pH or temperature relationships. One additional study (Kitamura 1990) was excluded from evaluation of the pH-TAN toxicity relationship because of missing information regarding the form of the ammonia addition, whether the ammonia concentrations were based on nominal total NH₄Cl additions or measured NH₄-N concentrations, and the temperature of the study. Also, because this article was written in Japanese, it was impossible to address some of these questions.

For evaluation of the pH-TAN acute toxicity relationship, the study by Wang et al. (2008) with Lampsilis demonstrated that ammonia toxicity to unionid mussels is pH dependent, and that the current pH-TAN acute toxicity relationship equation effectively represents the pH-TAN toxicity relationship for L. siliquoidea, as well as for other invertebrates, Potamopyrgus antipodarum (snail), Macrobrachium rosenbergii (freshwater shrimp), and H. azteca (amphipod), as tested by Hickey and Vickers (1994), Straus et al. (1991), and Borgmann and Borgmann (1997), respectively (p. 51 in the 2013 criteria document).

For the evaluation temperature-TAN acute toxicity relationship, the results revealed that the current temperature-TAN acute toxicity relationship equation effectively represents the temperature-TAN toxicity relationship for P. antipodarum, B. sowerbyi, and V. bengalensis as tested by Hickey and Vickers (1994) and Sarkar (1997), respectively.

No new (or other) data (including for freshwater fish) have been published (or made available) since 1999 to evaluate the pH- and temperature-TAN toxicity relationships; thus, it appears that the current equations are still applicable for the 2013 ammonia criteria update.

4.1.3. Peer Reviewer 3

Most of the chronic toxicity tests relied on by EPA for chronic freshwater criterion development are consistent with EPA's Guidelines and other water quality criteria developed. I am not convinced that the 28 day juvenile test data for the snails included in Table C should be used. Both of these tests used mixed-aged organisms, and in the case of the Ozark spring snail, organisms were also field-collected adding uncertainty in terms of their condition and acclimation to laboratory test conditions. In general, EPA has not relied on such tests for chronic criteria development, and I believe for good reason. I am not aware of any other relevant data that EPA did not already consider in this document.

4.1.3.1. EPA Response

Subsequent to the publication of the 2009 draft criteria, additional 28-day ammonia toxicity studies on the non-pulmonate snail species the pebblesnail (Fluminicola sp.) were performed to validate the data included in the 2009 draft ammonia criteria (Besser, 2011). EPA conducted a contractor-led external peer review of these more recent test data and the peer reviewers agreed that the toxicity data for the large juvenile snails were acceptable for quantitative use in the ammonia criteria derivation.

4.1.4. Peer Reviewer 4

Please see the response to question 1 above. That statement is also applicable to the chronic data base.

4.1.4.1. EPA Response

Thank you for your comment. EPA did conduct an updated literature search and included relevant chronic data published after the peer review to ensure the criteria document reflects best available current scientific information.

4.1.5. Peer Reviewer 5

The tests utilized for the criterion are scientifically defensible for such use and I'm not aware of other relevant data that were not used.

4.1.5.1. EPA Response

Noted, thank you.

4.1.6. Peer Reviewer 6

I believe that they are scientifically defensible for the use in the development of the chronic criterion. I am not aware of other relevant data.

4.1.6.1. EPA Response

Noted, thank you.

4.1.7. Peer Reviewer 7

Using the standards established by the Guidelines and the respective toxicity test methods the toxicity tests used to derive the chronic criterion are scientifically defensible. However, this does not mean that the data generated from these tests and the resulting test endpoints are scientifically defensible for the use. As stated before, there are standards that should be met to establish that the test results are of sufficient quality to be used in deriving water quality criteria. These standards include an appropriate concentration response curve and limits on variability within and between test treatments, replication of tests and use of reference toxicant tests to gauge organism sensitivity, as well as defensible control response across tests. One of the shortcomings of the chronic freshwater mussel method is that it does not include minimum organism size (weight, length) requirements that must be met at the end of the test in controls. It is unknown whether the growth expressed in the controls of these tests is acceptable in a natural environment where stress does not exist. The tests only can establish weights attainable in controls in the lab environment when organisms in the controls meet the minimum survival requirement. The organisms may be stressed in all test vessels, resulting in a bias of the test endpoint. The combination of treatment stressors (ammonia and vessels without sediment, for example) may produce a synergistic effect resulting in a lower ammonia IC20 for growth and survival than that resulting from a test which has only one stressor (ammonia). In this case, the control response cannot account for the synergism, even though the control is intended to account for all factors independent of the tested treatment.

I am not aware of any other relevant data that was not used.

4.1.7.1. EPA Response

See response 2.1.2.1.

4.2. Question 7: Is the freshwater chronic criterion scientifically defensible with mussels present and mussels absent?

4.2.1. Peer Reviewer 1

Yes. Similar distinctions for the presence or absence of certain sensitive organisms showing different sensitivities to ammonia have been used before. The 1999 ammonia criteria recommended different chronic criteria based on the presence or absence of early life stages of fish in the water body, setting a precedent of allowing different criteria for different situations. These different criteria were based on a similar magnitude of difference in sensitivity between the early life stages and the older life stages of fish (in the 1999 reassessment criteria) as is seen between the sensitivities of freshwater mussels and other species (in the 2009 reassessment). This policy of providing these separate criteria based on the presence absence of freshwater mussels will provide some flexibility in implementing the criteria in some instances and is reasonable.

Given the wide spread nature of ammonia in natural environments and the difficulties and expense imposed on municipalities in treating for ammonia in sewage treatment facilities, a significant lowering of the ammonia criteria is likely to have important consequences for municipalities. Providing these two criteria may help alleviate concerns that any additional restrictions on ammonia discharges are targeted to providing the needed level of protection for the waterbody, based on whether or not the more sensitive

mussels are present. Providing two EPA recommended criteria that already take these issues into account could help reduce the potential for requests for consideration of developing site-specific criteria based on the recalculation procedure, which can require significant resources to pursue.

Table C shows the four most sensitive genera and the calculation of the CCC for the two different criteria (with and without Unionid mussel data). The data under “excluding freshwater mussel data (family Unionidae)” includes the GMCV for the fingernail clam, *Musculium transversum* as the third most sensitive genera and this GMCV is used in the calculation of the CCC for “non-Unionid containing waters”. While it is true that *M. transversum* species is not in the family Unionidae, it is a bivalve mollusk.

In the recommendations for developing separate acute criteria, the calculation of the FAV and CMC as shown in Table B, the Asiatic clam was removed from the dataset to produce a criteria based on a “non-bivalve mollusk” dataset (as commented on under question # 2 above). There may be water bodies with physical habitats (such as some headwater streams) that are unsuitable for bivalve mollusks of any kind, not just Unionids. EPA should consider whether or not to base their recommendations regarding separate criteria based on removing data for Unionid genera only, or for all bivalve mollusks. The same approach should be used for both the acute and chronic criteria.

(Note: there is a small typo in Table C; the far right hand column should be labeled GMCV, rather than GMAY).

4.2.1.1. EPA Response

See response 2.2.1.1. [Note: All typos regarding use of the term GMCV have been fixed.]

4.2.2. Peer Reviewer 2

The freshwater chronic criterion based on a mussels present/absent approach maybe scientifically defensible – but will potentially result in very different approaches, even in regions of the United States where mussels are known to be absent or rare. In fact, the primary question will become “what level of data is required” to determine whether the habitat is or is not suitable to support mussels and/or whether mussels have ever occurred within the site historically. Guidance on such data needs questions would be useful to include in the criteria document to support a with/without mussel criterion.

Additionally, even when the “mussel present” option is applied, a whole other series of issues arise, because there is often a poor correlation between the water column NH₃ concentration and the pore water NH₃ concentration where most juvenile mussels reside. Site-specific sediment conditions can greatly influence the availability of NH₃, and in the case of the upper Mississippi River, as much as 6-30 times greater than the surface waters (Bartsch et al. 2003).

In such cases, criteria developed without considering the complete affects of sediment on the accumulation of NH₃, or other confounding affects on mussel communities (i.e., the fact that mussel beds can self-produce nutrients, such as ammonia) (Dankers and Koelemaj 1989), may greatly underestimate, or even overestimate, a criterion that is protective of mussels. In addition, this also brings up the question of whether the use of sediment is appropriate in chronic toxicity tests designed to be protective of aquatic life in the water column, and what endpoint should be evaluated using such methodologies.

4.2.2.1. EPA Response

See response 2.2.1.1.

4.2.3. Peer Reviewer 3

I don't think it is. See answer to question 2. The chronic value used for *L. fasciola* in Table 4 (0.39 mg N/L), and which must have been used in the GMCV calculation in Table C, doesn't appear to make sense given an IC₂₅ at the same pH and temperature reported of 0.38 mg N/L on p. 20 and the IC₂₅ based on

the actual data (pH = 8.2 and 20°C) of 0.39 mg N/L. The IC or EC20 should be lower than the IC or EC25. The IC pin value of 0.23 mg N/L appears more in line with the data for this species. EPA should at least defend why 0.39 mg N/L is a more appropriate value than 0.23 mg N/L for this species. In addition, wouldn't it be more consistent to use the ICp value for both *Lampsilis* species because the *L. siliquoidea* chronic value could only be calculated using the ICpin method?

4.2.3.1. EPA Response

EPA did not use the ICpin program [Linear Interpolation Method for Sublethal Toxicity: The Inhibition Concentration Approach (ICpin, Version 2.0, 06/03)] to develop the 2013 final ammonia criteria. The TRAP program (EPA's Statistical Program: Toxicity Relationship Analysis Program, Version 1.21)) was used to calculate EC20 values/GMCVs to develop the chronic aquatic life criterion for ammonia.

4.2.4. Peer Reviewer 4

The chronic criterion is defensible, provided that suitable guidelines are developed to differentiate environments where mussels are present from those where they are absent. Please see my response to question 2 above (see Comment 2.2.4 on p. 10).

4.2.4.1. EPA Response

The criteria are no longer bifurcated. See EPA response to peer reviewer 1 comment (Response 2.2.1.1 on p. 8)

4.2.5. Peer Reviewer 5

The criterion does not seem defensible for locations where mussels are present as would relate to considerations under the Endangered Species Act, but would likely be defensible under considerations with the Clean Water Act. I pose this statement as an example of the difficulty in relying on scientific defensibility in the face of a very broad range of considerations in risk and decision management that may rely on these estimates.

4.2.5.1. EPA Response

See response 2.2.1.1.

4.2.6. Peer Reviewer 6

Similar to the acute criterion, I believe it is scientifically defensible to develop the criterion with mussels present and mussels absent. I do believe that the EPA should have a cautionary note as in the acute criterion as to when and how to determine if mussels are present or absent.

4.2.6.1. EPA Response

See response 2.2.1.1.

4.2.7. Peer Reviewer 7

As stated in question #2, EPA should determine whether simply the presence of mussels is important or the presence of specific species is important. The first step in determining whether mussels are present or absent is to define a standard for the conditions describing that a population exists at a site. This is not a straight-forward exercise and can be quite difficult as seen in various site-specific criteria projects around the country. Guidance must be provided as to what parameters and measures of those parameters are to be used to define presence/absence of a population. Quantitative evidence of reproduction would be required at a minimum, as well as the presence of various developmental stages. The normal abiotic conditions documented for a species should also be compared to that of a particular surface water to

determine whether a conclusion of presence or absence is supported. Certainly the availability of host species could influence the determination of absence/presence.

Once absence/presence is established EPA should only use mussel data to develop water quality criteria for waters where mussels are deemed present. EPA should ensure that conditions of the site where absence has been concluded have not been anthropogenically changed to a state where mussels cannot reproduce and biological records have documented the presence of previous populations. Historic records of water quality and biological surveys would help support this determination. It is clear that EPA and the states will share the burden of the presence/absence determination and that there will be differences of opinion as to whether, when absence seems apparent, conditions changed due to human activity that could have been reasonably precluded or the species simply never successfully established a population at the site due to species-specific factors. Given that the chronic sensitivity of mussel species drives the CCC approximately 44% lower than when mussel data is not used the management actions and costs based on criteria without mussel data could be significantly different than when mussel data is used to develop the CCC.

The presence/absence of species is usually a question that is asked when criteria are recalculated or a site-specific criterion is being developed. In these cases the issue of removing species from a database is addressed by determining whether data for other species with similar taxonomy is available. The taxonomy and reproductive cycle of the freshwater mussels is so unique that this will not be possible. Therefore a decision must be made as to whether to apply a CCC with or without mussel data. Given that there is so little chronic data available for these mussels and that there are outstanding issues that must still be addressed (use of shell measurements to represent growth, tests without sediment, irregular concentration response curves, etc.) a CCC with mussel data will result in water quality standards with uncertain consequences. This uncertainty demands that criteria without mussel data be available for states to use when mussels are not resident (reproducing population found year to year). At this point in time a CCC without mussel data would be more scientifically defensible than one with mussel data.

4.2.7.1. EPA Response

See response 2.2.1.1.

5. USE OF 28-DAY JUVENILE TEST DATA: (see Appendix C IN 2009 DRAFT CRITERIA UPDATE):

Water quality criteria for the protection of aquatic life are derived using toxicity endpoints that relate to population level impacts. In general, these endpoints relate to survival, growth and/or reproduction. The 28-day test with juvenile mussels, while similar in duration to a standard chronic test, is not technically an early-life stage test according to the 1985 Guidelines for Aquatic Life Criteria, as much of the early development will have already occurred.

5.1. Question 8: Given that the juvenile life stage of freshwater mussels is relatively long (2-6 years) are 28-day exposure tests with juvenile mussels scientifically defensible as "chronic" test data for criteria development?

5.1.1. Peer Reviewer 1

I do not believe that growth measured during a 28-day test can be considered a true measure of a chronic effect as intended by the requirements in the 1985 Guidelines for Aquatic Life Criteria, such that they can be related to significant population impacts. Given the life span of the organism and the time needed to reach maturity, a 20% difference in growth observed over 28 days may be reversible and may not be relevant on a population level and should not be used to derive criteria.

5.1.1.1. EPA Response

*Growth data from 28-day tests with juvenile unionid mussels presented in the Wang et al. studies (2007, 2011) were not used to derive the 2013 final chronic criterion for ammonia. EPA's decision not to use the growth data was due to uncertainty in the test methods for assessing the growth endpoint and the need, as stated by the authors, for additional research "to optimize feeding conditions, to conduct longer-term exposures (e.g., 90 d), and to compare growth effect to potential reproductive effect in partial life-cycle exposure" (Wang et al. 2011). The growth endpoint showed a high degree of variability, and the test methods for assessing growth, based on substrate or water-only exposures, are currently being evaluated. For these reasons, the survival data but not the growth data for 28-day juvenile mussels were used in the calculation of the 2013 chronic ammonia criterion. Appendix G in Aquatic Life Ambient Water Quality Criteria for Ammonia – Freshwater (2013) provides the TRAP-estimated EC20s for survival for rainbow mussel and both *Lampsilis* species, and a comparison to the growth of fatmucket mussel from 28-day tests reported by Wang et al. (2007a, 2011), which shows the uncertainty in the concentration-response relationship for the growth endpoint. The growth data are cited as "Other Chronic Data" in Table C of EPA's final Aquatic Life Ambient Water Quality Criteria for Ammonia – Freshwater (2013).*

EPA has reconsidered the previous position and rationale for excluding the 28-day ASTM method for juvenile mussels as a chronic test. For juvenile unionid mussels, a life stage that lasts 2-6 years or longer, 28-day tests are relevant to consider and parallel somewhat the use of 28-day test data for salmonids with similar or longer life spans in chronic effects assessments (see response to Peer Reviewer 2 comment 5.1.2 below).

5.1.2. Peer Reviewer 2

Water quality criteria for the protection of aquatic life are derived using toxicity endpoints that relate to population level impacts. In general, these endpoints relate to survival, growth and/or reproduction. I would agree that the 28-day test with juvenile mussels, while similar in duration to a standard chronic test for many other groups of organisms, is not technically an early-life stage test according to the 1985 Guidelines for Aquatic Life Criteria, and as such, should perhaps not be included in the chronic database.

If this is carried forward, I would support the approach of the 28-day exposure tests being considered as "other data" – if the test are acceptable in all other respects. Reasons for not including in chronic criteria calculations could follow and build on the arguments presented in the position paper:

- First, it is not representative. If the duration of the life stage is 2-6 years for juvenile life stage of fresh mussels, 10% of this amount of time would be 73-219 days. Thus, a longer exposure (i.e., 90-day) would be more representative for a "chronic" study for juvenile fresh water mussels than 28-day tests.
- Second, extending the exposure time would allow the tested animals to adapt better to laboratory conditions, thus reducing the stress and decrease the risk of having test condition-biased results.
- Third, although it will be more challenging to keep all the water quality parameters under control in a longer exposure it could also provide other avenues for future research on juvenile mussels (i.e., if captivity conditions and pollutant exposure would enhance growth over longer periods of time, if they may need additional food source, among others)

One possible conflicting issue is that other species used in EPA criteria development have "allowable" chronic test conditions that also represent a small portion of their total life span – e.g., 60-day post hatch ELS for salmonids, which can live for years. The difference, conceptually, is that the 28-day test for mussels does not include the earliest life stage prior to juvenile testing commences – i.e., it does not start with egg, then hatch, then juvenile exposure more typical of an ELS for fish. If we can conceptually overcome this "weakness" for a 28-day juvenile mussel test, then I suppose growth-effects data in a 28-

day juvenile mussel test could be at least evaluated in the context of other reported data – and if comparable to effects data from longer-duration testing, then perhaps it could be used in some form for criteria calculations.

5.1.2.1. EPA Response

Considering the decline of freshwater mussels in North America with numerous species listed as Federally endangered or threatened it was important to consider available data on freshwater mussels that would be relevant for chronic effects in developing the chronic criterion for ammonia. See response 5.1.1.1.

5.1.3. Peer Reviewer 3

The question posed here by EPA implies that a 28-d test may not be sufficiently long for freshwater mussels that live 2-6 years (actually, many unionid species live far longer than that—20 years is not uncommon). If this is the case, I fail to see how EPA justifies use of 28-d tests for fish such as salmonids that live 3-5 years or sturgeon that live > 30 years. I realize there's McKim's 1977 paper but that was based on outdated test methods and has questionable relevance to data collected in more recent years.

The discussion concerning the usability of 28-d mollusk survival and growth tests in chronic criteria development (pages 15-16 of the draft document) appears to have contradicting statements and questionable rationale in my view. The discussion begins with the statement that 28-d tests do not qualify as life-cycle, partial life cycle, or early life stage tests as defined in EPA's Guidelines. According to those guidelines, 28-d tests can only be used for fish because it has not been demonstrated that a 28-d exposure is a reasonable predictor of invertebrate sublethal chronic effects (defined as at least 90d for vertebrates and presumably for invertebrates such as mussels which live for several years). However, the draft document goes on to say that 28-d survival information could be used to inform the chronic criterion because obviously, lethality is not reversible. Yet in footnote #2 on page 15 of the document, as well as in Appendix C, uncertainties regarding the 28-d mussel test method are discussed which include optimal quality and quantity of food, exposure apparatus, etc. needed to sustain juvenile mussels in good health and maintain adequate growth throughout the test. Based on my limited experience testing bivalves, I believe it may not be as challenging as one might expect to keep control juvenile mussels alive for 28-d given adequate oxygen and clean water but it may be quite another thing to document that organism condition (or alternatively, susceptibility to stress) is truly adequate. Therefore, the fact that juvenile mussel survival is lower with exposure to certain levels of ammonia may be due to having an already stressed population due to inadequate food or other factors.

A similar finding is now being made by EPA after reviewing chronic *Hyaella* chronic test data (reported to be one of the most sensitive species in the 1999 ammonia criteria document). In that case, control survival was apparently satisfactory but test organisms were in fact ultrasensitive to the chemical due to suboptimal test conditions.

A second point I would raise regarding the chronic mussel tests is that they were only conducted once (based on published information) for a given species in one laboratory. While this in itself is not a reason to exclude test data, it is clear that there is considerable variability in acute endpoints among juvenile mussel tests using the same species and conditions (even within the same lab as mentioned previously), which should apply to chronic survival with ammonia as well. Given that certain mollusks appear to be far more sensitive to ammonia than any species tested thus far, there should be some replication of these results to be sure we're not going to find out that EPA needs to do yet another reassessment because of issues similar to the ones they are dealing with now for *Hyaella*.

It is important to note that while interlaboratory testing was conducted for acute juvenile tests (Wang et al. 2007), similar interlab testing has not been conducted (or at least published) to my knowledge for 28-d mussel chronic tests. In addition, one should view the acute interlaboratory results cautiously because, as

mentioned previously, a single culture and water source was used by all laboratories in that study. As Wang et al. (2007) point out: “However, the present study used the same batch of test organisms and the same dilution water, which were different from the referenced inter-laboratory studies and might have contributed to the lower variability in test results. The present study was designed to determine the inherent variability in the test. Higher variability would be likely if the inter-laboratory tests were conducted with different sources of dilution water and organisms (e.g., from different populations or watersheds). Therefore, additional study is needed to further characterize potential variability associated with the newly developed ASTM standard methods for conducting acute toxicity tests with early life stages of freshwater mussels.” Therefore, the CVs reported by Wang et al. (2007) are likely to be underestimates of true test method variability. This is borne out by the acute data generated for the same species by different labs as noted previously in my comments

5.1.3.1. EPA Response

See response 5.1.1.1.

5.1.4. Peer Reviewer 4

This is well covered in Appendix C, and the arguments are convincing. I agree that the 28-d test does not qualify as a chronic test, since it is neither a life-cycle test nor a partial life-cycle test. I also agree that a concentration causing a greater than 20% reduction in survival could be used as an upper limit for the Species Mean Chronic Value, assuming the test met the requirement for good lab practice.

5.1.4.1. EPA Response

See response 5.1.1.1.

5.1.5. Peer Reviewer 5

Research that validates effective endpoints related to the range of pollutant effects in freshwater mussels have shown that 28-30 -day exposure tests in both field and laboratory settings can be included to predict long term impacts related to biomarker responses, bioaccumulation, and even comparisons to conventional “chronic” test data. The 28-day test exposure period will continue to be temporal as relates to the life cycle, but can be relational to long-term considerations and offer critical information when combined with multiple lines of evidence. Inclusion of more than one life stage response also advances mussel toxicity assessments, rather than providing semantic arguments over longevity for a life stage and whether the exposure period qualifies as “chronic”. Data from the 28-day test can be defensible when supported by “other” data. Other data should also be scientifically defensible.

5.1.5.1. EPA Response

See response 5.1.1.1.

5.1.6. Peer Reviewer 6

I do not believe that 28 day mussel studies should be considered as chronic test data for the national database as described in the position statement and supporting documentation. I believe this is very well thought out and applaud the developers.

5.1.6.1. EPA Response

See response 5.1.1.1.

5.1.7. Peer Reviewer 7

This question is not only a function of the duration of exposure in the test but whether the test meets the requirements of the Guidelines. Clearly the test does not meet the Guidelines requirements because it

does not address either the entire life cycle or, for partial life-cycle tests, the appropriate life stages and test endpoints defined by the Guidelines. The reliability of this test is also clouded by the apparent impact of testing juveniles in the absence of sediment and the use of shell length as a chronic test endpoint for freshwater mussels (see comments below). Perhaps the Guidelines require updating (has not been formally revised since its first release in 1985), but until this is done the 28-day juvenile test does not meet the standard for scientific defensibility that EPA has established.

One might suggest using the data from these tests as an estimate of freshwater mussel chronic toxicity because the tests did not include the glochidia life stage and it is more appropriate than not using the data for criteria development. Unfortunately there is enough concern regarding the reliability of this test even to accurately represent the sensitivity of juveniles that it would not be appropriate to use the data in any water quality criteria capacity.

Independent of the issues raised above, it would seem that a 28-day exposure test with juvenile mussels would be representative of the sensitivity of juveniles. However, chronic tests used for developing a CCC are intended to evaluate or extrapolate the sensitivity of a population to a stressor. This requires reliable knowledge of the sensitivity of the most sensitive life stages of that population. It is unknown how the sensitivity of glochidia compares to juveniles because of the issues associated with both acute and chronic test methods. Without this knowledge and reasonable certainty it cannot be determined that the 28-day juvenile test with freshwater mussels is scientifically defensible relative to development of a CCC.

5.1.7.1. EPA Response

See response 5.1.1.1.

Based on results from acute ammonia toxicity tests, glochidia were not consistently more sensitive than the juvenile mussels thus data for both glochidia and juveniles were used in the calculation of the acute criteria. Due to the limited duration of the glochidia life stage (viable up to 10 days depending on species) a chronic toxicity test with glochidia is not possible.

5.2. Question 9: Should toxicity studies on the growth rates of mussel shells during 28-day tests be considered quantitatively when developing water quality criteria?

5.2.1. Peer Reviewer 1

No. I do not believe that differences in growth rates measured over a 28-day period should be considered to provide acceptable quantitative chronic endpoint data for freshwater mussels that should be used to calculate chronic criteria for ammonia. Chronic criteria for ammonia must be based on clearly recognized adverse effects in order to be acceptable. Reductions in survival are irreversible and are clearly adverse effects, while slight reductions in growth may not be significant to the population, or the organisms may be able to compensate over time and reverse the difference in growth. These type tests do not meet the requirements of the EPA Guidelines for use as a quantitative value in driving a chronic criterion. Using such data to derive chronic criteria, especially if it significantly influences the final criteria would likely increase uncertainty about the final criteria.

5.2.1.1. EPA Response

See response 5.1.1.1.

5.2.2. Peer Reviewer 2

As described in the previous question, the studies on growth rates of mussels shells during 28-day tests should only be used as information to better evaluate appropriate longer-term chronic studies (i.e., 90-day

studies). The results from longer chronic tests could certainly be used quantitatively for the developing the water quality criteria. Unless 28-day test data provide similar endpoints to long-term testing, it would be necessary to keep the results separate and potentially have two conflicting sets of data (i.e., 28-day and 90-day data) – which will only make the calculations of the new criteria more complicated, and can lead to confusion when trying to implement. And I have to say that ammonia criteria development is confusing enough as is.

5.2.2.1. EPA Response

See response 5.1.1.1.

5.2.3. Peer Reviewer 3

I don't believe there is sufficient information or adequate test method development and standardization at this time to be able to use growth rates of mussel shells in 28-d tests for criteria development. Growth rates, based on shell measurements, are not sufficiently large over a 28-d period to have much confidence in the results. Based on the studies reported in this draft, as well as others of which I'm aware, shell length is too coarse a measurement to be used as an indicator of mussel growth over such a short time period. Furthermore, it is generally not known what the expected growth rate should be for a given species under these test conditions. As I indicated earlier, without having such information, I don't believe that survival data from such tests can be used with confidence as well.

5.2.3.1. EPA Response

See response 5.1.1.1.

5.2.4. Peer Reviewer 4

The approach taken in the document (Appendix C), that of not using data for growth of organisms in non-life-cycle tests with durations of less than 90 days in the derivation of a Species Mean Chronic Value (SMCV), is in my opinion perfectly appropriate. The justification given in paragraph 5 of Appendix C centers on two points: lack of direct comparisons of growth between shorter tests and life-cycle tests and uncertainties around the relationship between short-term growth reductions and growth effects over a full life cycle. Both are valid reasons to exclude the data from calculations of SMCV.

5.2.4.1. EPA Response

See response 5.1.1.1.

5.2.5. Peer Reviewer 5

Growth rates of mussels during the 28-day tests should not be considered quantitatively until more information is available on the validity of this test. Variability is still attributable to diet and holding conditions and has been cited broadly.

5.2.5.1. EPA Response

See response 5.1.1.1.

5.2.6. Peer Reviewer 6

The problem with using shell growth rates, it doesn't necessarily link to a smaller organism being any less viable or having less reproductive benefits. Since mussels are not a commercially or recreational important organism, I believe that the use of growth rates are not valid as compared to a fish species or commercially important bivalve.

5.2.6.1. EPA Response

See response 5.1.1.1.

5.2.7. Peer Reviewer 7

Review of Wang et al. (2007) for both the copper and ammonia 28-day tests shows that there is a very small difference between the length of the shells at the beginning and at the end of the test based on control responses. In fact it appears that some of the tests show no statistical difference in shell length between the beginning and end of a test at $\alpha=0.05$ and assuming a normal distribution. This does not support the use of shell length to represent growth in the 28-day mussel tests. Further this paper shows that using shell length to represent growth does not provide reliable dose response curves in most cases. Although the authors only conducted 21-day tests both Bringolf et al. (2007a) and Bringolf et al. (2007b) found that the concentration response curves for different chemicals were not what one would expect for a toxicant. Although there may be issues with organisms benefitting from certain concentrations of chemicals, this fact only complicates the interpretation of this data and increases the uncertainty of using toxicity test results based on shell length.

5.2.7.1. EPA Response

See response 5.1.1.1.

5.3. Question 10a: Regarding the position statement and rationale on use of juvenile mussel growth data – Is it scientifically defensible to include the juvenile growth data from a 28-day exposure period as “other data” that might influence the criteria however not be used directly in the derivation of the criteria value?

5.3.1. Peer Reviewer 1

Yes. Although too many questions surround the use of growth effects during a 28 day test in relation to long lived species such as freshwater mussels to consider using differences in growth as data that directly influence the calculation of the final chronic criteria, these data still deserve consideration and inclusion in the EPA criteria document. These data do provide useful information to help judge the potential for “other” effects that have a potential for concern. Including such data under the category of “other data” can provide additional information that can lend confidence that the criteria based on survival can provide an adequate amount of protection. The magnitude of difference in growth may make a difference; a 20 to 25 % reduction in growth may not be of much ecological significance, especially if there is a thought that the effect may be short term and reversible, while a > 75% reduction in growth may be considered more of a concern. The magnitude of difference between the effects on survival and the effect on growth can also provide useful information. Altogether, I see value in presenting these growth data under the classification of “other data” in the criteria document.

5.3.1.1. EPA Response

See response 5.1.1.1.

5.3.2. Peer Reviewer 2

My understanding of data included as “other data” in the criteria document is that these data are provided just as a reference to make the interested parties know that these data exist. They can certainly be presented that way and not considered in the calculations of the chronic criteria for the reasons outlined in the previous two questions

5.3.2.1. EPA Response

See response 5.1.1.1.

5.3.3. Peer Reviewer 3

I agree with EPA's position that mussel growth data should be categorized as "other data" until it is demonstrated that there is a repeatable, defensible methodology for generating adequate growth data and that the growth endpoint calculated is ecologically meaningful. For this reason, I don't think it matters whether impaired growth observed in a 28-d test is indicative of effects at longer exposures or not. However, I would recommend that EPA not consider 28-d mussel growth as a valid endpoint based on methodological and perhaps biological uncertainties, as I mentioned under (9) above, rather than because reduced growth rate is theoretically a reversible effect. I see no reason why reduced growth would not be an even bigger effect with increased exposure duration to a toxicant.

5.3.3.1. EPA Response

See response 5.1.1.1.

5.3.4. Peer Reviewer 4

With respect to the first point above, inclusion of the data as "other data" is perfectly appropriate. This is consistent with the practice for short-term growth results with other species, and I see no reason why mussel data should be treated differently.

5.3.4.1. EPA Response

See response 5.1.1.1.

5.3.5. Peer Reviewer 5

Growth should not be considered as defensible for the above mentioned reasons relating to lack of reliable test conditions, but can be used as supportive information that might influence the criteria. Studies that include recovery and or stability in growth or even degrowth with freshwater bivalves, have produced more reliable estimates of pollutant effects and could therefore be used to influence the criteria when included with broader scale studies.

5.3.5.1. EPA Response

See response 5.1.1.1.

5.3.6. Peer Reviewer 6

I believe that it should remain in the other data section so the relevant data can be used in site specific evaluations.

5.3.6.1. EPA Response

See response 5.1.1.1.

5.3.7. Peer Reviewer 7

Based on the uncertainty associated with testing juveniles without sediment and the lack of reliability in the shell length endpoint it is not appropriate to use the chronic data in any way associated with development of a CCC.

5.3.7.1. EPA Response

See response 5.1.1.1.

5.4. Question 10b: Regarding the position statement and rationale on use of juvenile mussel growth data –Should the statement also consider impaired growth of mussels which were affected at a 28-day exposure could as likely continue to decline in longer exposures as another potential outcome (i.e., the chance they could recover or stabilize is one potential outcome only)?

5.4.1. Peer Reviewer 1

I would agree that this statement could also be true. We simply do not know whether or not an effect on growth observed over a 28-day exposure can be considered reversible or indicative of a permanent effect. As discussed above, I believe the magnitude in difference in the growth and the magnitude of difference between the survival end point versus the growth endpoint could influence the level of concern regarding such data.

5.4.1.1. EPA Response

See response 5.1.1.1.

5.4.2. Peer Reviewer 2

As noted in the position paper, it is not possible with the available data to conclude that impaired growth after 28-day exposure will necessarily lead to continued reductions in growth over time. The position paper rightly points out the possibility of organisms to “recover” from potentially short term stress and show no further decline. In fact, these organisms may exhibit normal growth during a more relevant long-term test. However, given the lack of information available to compare short-term and long-term toxicity studies in fresh water mussels, it would be advisable to note, as appropriate, that any of the three outcomes: continued decline, stabilize, or recovery is possible. In addition, the propensity for any of these three outcomes will most likely depend upon the particular mussel species and the stage of development used in the toxicity tests.

5.4.2.1. EPA Response

See response 5.1.1.1.

5.4.3. Peer Reviewer 3

(see above response for 10a)

5.4.4. Peer Reviewer 4

With respect to the second question, this is a statistically valid point, but it does not reduce our level of uncertainty around the actual meaning of the data. The issue will only be resolved when someone undertakes 90-day tests.

5.4.4.1. EPA Response

See response 5.1.1.1.

5.4.5. Peer Reviewer 5

As stated, the above consideration seems very confusing and should be clarified to relate to the defensibility as considered with “other data”.

5.4.5.1. EPA Response

See response 5.1.1.1.

5.4.6. Peer Reviewer 6

I don't believe that data supports this and such a statement should not be made.

5.4.6.1. EPA Response

See response 5.1.1.1.

5.4.7. Peer Reviewer 7

The comments provided above on the reliability of the shell length endpoint for chronic tests of freshwater mussels indicate that it is not a reliable endpoint to represent the chronic sensitivity of these organisms. Without more data, it is not defensible to speculate on the outcome of tests with longer duration when the reliability of the endpoint being used is in question.

5.4.7.1. EPA Response

See response 5.1.1.1.

5.5. Question 11: The values of the acute and chronic ammonia criteria have a strong dependence on pH. Juvenile and adult mussels, as sediment-dwelling organisms, inhabit a medium that may have vertical pH gradients, thereby creating some uncertainty about the appropriate pH to assign as their exposure conditions. For applying a criterion protecting mussels, do you have suggestions on how states and EPA might determine the pH applicable to the sediment micro-environment to which mussels are typically exposed?

5.5.1. Peer Reviewer 1

The toxicity data currently available for freshwater mussels are for glochidia and juveniles. The exposure to unencysted glochidia will be primarily via water exposure, and once encysted on a host fish they will be in a less exposed condition. Juvenile mussels will also be initially exposed to ammonia in the water column following dropping off (excystment) from the host fish. As the juvenile mussels are deposited onto the substrate they may remain on the surface or may become situated within a few centimeters of the surface. They will need to maintain a position on or near the surface in order to ensure access to well-oxygenated water as well as filterable food. In all cases, the exposure to ammonia will be over the gills via water; either surface water or interstitial water and the water quality criteria can be adjusted to account for any differences in pH between the water column, the surface-sediment interface or pore water at a specific depth in the sediment.

For normal use in establishing permit limits, or assessment of waterbodies; the criteria should be applied to the water column. If there is a special concern at a site, the option of monitoring the pH of interstitial water can be investigated, and the pH characteristic of the interstitial water can be used to assess the criteria. Care should be taken to determine the pH in the interstitial water at depths where mussels are likely to be found at the site.

5.5.1.1. EPA Response

Thank you for your comment. EPA sought this advice to support critical evaluation of new or improved test methodology(ies) for mussels, and for potential future questions or considerations regarding criteria implementation. The Agency's 2013 final national ammonia criteria recommendations are designed to be protective of aquatic life exposed to water via the water column, including for benthic macroinvertebrates such as mollusks. The national criteria are presented at a specific pH and temperature which are meant to be representative of water quality conditions within the water column. All of the data are normalized

to the same pH and temperature. Extrapolation tables are provided in the 2013 criteria document for determining the applicable criteria at other pH and temperatures within the water column.

5.5.2. Peer Reviewer 2

Given that juvenile mussels of most species reside completely in the sediment, filter pore water, and feed on the sediment, I would suggest one approach might be determining a mean burial depth for the juveniles of the selected mussel species within the 2009 data set. In laboratory tests, juvenile *Villosa iris* are reported to bury <1cm, yet were not exposed to the overlying water (Yeager et al. 1994). Perhaps it would be useful to use this burial depth and evaluate the average pH conditions within a burial zone, per se. In addition, it is likely any measures of pH in the mussel micro-environment (i.e., taking measures of water quality over a mussel bed) may also show elevated ammonia levels simply from the metabolic activity of those organisms, themselves. I can't help but wonder how to determine attainment of "safe" ammonia levels in future field monitoring, given the potential for organism-generated ammonia values.

5.5.2.1. EPA Response

See response 5.5.1.1.

5.5.3. Peer Reviewer 3

I am not so sure that applying ammonia criteria to juvenile mussels at least, and possibly adults as well, should consider sediment pH at all for several reasons. First, larger juvenile and adult mussels are reported to be exposed to waterborne contaminants via siphoning the overlying water and not as much via sediment ingestion or pore water exposure during their growth season (Strayer et al. 2005). This is particularly true during times when water temperature exceeds 50° F (Watters et al. 2001; Schwal and Pusch 2007). In fact, many mussel sampling protocols specify a minimum water temperature of 50-60° F because mussels are near or on the surface under those conditions and more readily seen (e.g., Pennsylvania DEP and Army Corps sampling protocol in the Allegheny River, PA). During this time (typically spring, summer, and fall in most waterbodies of the temperate U.S.), mussels are actively feeding via siphoning the water column and sediment pH is probably not a factor affecting water column ammonia toxicity. Rather it is the overlying water pH that is probably more important.

Second, many researchers report that sediment pH and ammonia can vary substantially from one location to another within even small areas and may be dependent on very complex redox gradients that are difficult to measure much less predict (Burton 1992; Stemmer et al. 1990; Sarda and Burton 1995). It is no easy matter to identify an appropriate sediment pH that should be used to apply an ambient water ammonia criterion.

Third, many of the mussel species tested thus far, or that are species of concern, inhabit well-aerated sediments that often have either fairly coarse particle size (gravels – cobbles) or sand, both of which have interstitial water closely resembling the water quality conditions of the overlying water (Salmon and Green 1983; Neves and Widlak 1988; Way et al. 1990). In these cases one would not expect a large difference between sediment and overlying water pH.

Fourth, it seems premature to consider complex pH adjustments based on sediments when it hasn't been demonstrated that juvenile or adult mussels are as sensitive to sediment ammonia as they appear to be to water column ammonia in lab tests. If indeed sediment pH needs to be considered, then it seems likely that mussels are not being exposed to surface water (including surface water ammonia) but rather sediment and interstitial ammonia. If this is the case, there is no need to adjust water column ammonia criteria based on sediment pH.

Finally, basing a water column criterion on sediment water characteristics represents a major departure for EPA in terms of criteria development. While I would encourage the Agency to consider more holistic criteria that take into account multiple media (e.g., sediment and water), I don't believe there is sufficient

information at this time to warrant altering a water quality criterion, which was derived using water exposure test methods exclusively, due to sediment characteristics.

5.5.3.1. EPA Response

See response 5.5.1.1.

5.5.4. Peer Reviewer 4

While mussels live in sediment, the pH of importance from an ammonia criteria perspective is the pH of the respiratory environment, which in this case is the pH at the sediment-water interface. The sediment pH is of importance only to the extent that it influences the pH of the water at the interface. Remember these are water quality criteria, not sediment quality criteria. I have no suggestion as to the appropriate method for determining the pH at the interface. The methodology, more in the realm of environmental chemistry, falls outside of my area of expertise.

5.5.4.1. EPA Response

See response 5.5.1.1.

5.5.5. Peer Reviewer 5

Even if an adequate amount of information was available for any benthic invertebrates for chronic toxicity related more specifically to the generic relationship of ammonia toxicity to pH, larger issues challenge this consideration with the reassessment for changes related to pH shifts, which may increase when phase waters are isolated for testing. These relationships have only recently been examined for juvenile freshwater mussels, without sufficient detail furnished as to the pH adjustment used with diluter systems in those tests or their effect upon alkalinity changes measured during the assays (Wang et al., 2007). Although these tests seem suitably performed to establish the influence of pH on the acute toxicity of ammonia and even the pertinence of normalizing ammonia toxicity data for mussels to a common pH end point (total ammonia nitrogen at pH 8.0), information is still lacking for the sediment relations specific to mussel responses. As relates to this consideration, the bulk of the studies that have examined sediment and phase water relationships to date rely heavily on marine field monitoring, and not standardized laboratory testing or field monitoring with freshwaters (Salazar and Salazar, 2007). Any scientifically defensible way to account for these differences would greatly benefit from studies with comparable methods examining habitat preferences, feeding, and mobility of a sensitive mussel life stage (juvenile) among sediment compartments, since the synthesis of ammonia toxicity data confirmed its status as a sensitive species. Data from these studies would also be needed to generate pertinent modeling of the ammonia fluxes from sediment and the concentration profiles in the phases that would have consequence to mussel impact.

5.5.5.1. EPA Response

See response 5.5.1.1.

5.5.6. Peer Reviewer 6

It is much easier to do this for water column organisms. I believe that the agency will open themselves up for tougher battle if they attempt to make this issue even more complex than it already is. I believe that noting that this is the case and stating that pH dependence vs. the criteria as it relates to mussels was not developed at this time given the uncertainty about what the appropriate pH assignment should be.

I also argue whether pH in sediment is a relevant question with regards to criteria since the management tool for water quality is going to be primarily the NPDES permit. How will anyone be able to control sediment pH to manage it? I am not sure what this does for development of a defensible criterion.

5.5.6.1. EPA Response

See response 5.5.1.1.

5.5.7. Peer Reviewer 7

It will be necessary for states and EPA to directly measure the pH of the environment where each life stage exists. For juvenile and adult mussels this will require measurements of the sediment in each location where the criteria are applied as water quality standards. Guidance from EPA on how to appropriately represent the pH of sediments where ammonia criteria are applied to protect freshwater mussels will be necessary. Measurements in the field will also be necessary because the pH of a discharge regulated by a NPDES permit cannot be used to estimate the pH of the sediment. The pH of the sediment can only be assessed through direct measurement. This may be done in situ or by rapidly measuring the pH of relevant sediments collected with sediment grab or core equipment.

5.5.7.1. EPA Response

See response 5.5.1.1.

5.6. Question 12: In general, should the criteria include a consideration for the potential pH difference between sediment and the water? If so, what is the most scientifically defensible way to account for these differences when deriving protective water quality criteria?

5.6.1. Peer Reviewer 1

The potential differences in pH between water and sediment is not so much directly related to the criteria development, as it is an issue related to how and where the criteria may be applied. The water quality criteria should provide recommendations for protective criteria based on ammonia concentrations in water and adjusted for pH and temperature. On a site-specific basis, if in sediment there are significant differences between pH in the interstitial pore water in sediment and the overlying surface water, this can be assessed by site-specific measurement of pH levels in the sediment (at depths where juvenile mussels are likely to be found) and then adjusting the criteria to those pH conditions. Another concern is whether or not the pH at the appropriate depth is relatively constant. These issues will likely vary considerably depending on the characteristics of the sediment, and perhaps seasonally. The criteria can be adjusted to the pH in the interstitial water to assess the potential for adverse effects at the site.

5.6.1.1. EPA Response

See response 5.5.1.1.

5.6.2. Peer Reviewer 2

Because adult mussels siphon directly from the water column, there is no need for a pH differential calculation. However, to be protective of the juvenile life stage which typically filters pore water, there may be a need for differential pH measurements. But how does one account for all of the biotic and abiotic factors that affect pH in the sediment? Because site-specific sediment conditions greatly influence the pH levels, adjusting the pH should only be an option for site-specific calculations (or considered as “other data”) rather than within the national water quality criteria.

5.6.2.1. EPA Response

See response 5.5.1.1.

5.6.3. Peer Reviewer 3

I don't think criteria should consider sediment pH or the difference between sediment and water pH in deriving water quality criteria. Rather, EPA should establish sediment ammonia criteria, using toxicity test data for mussels and other benthic species (e.g., *Hyalella*, *Hexagenia*, etc.) exposed to sediments with ammonia.

5.6.3.1. EPA Response

See response 5.5.1.1.

5.6.4. Peer Reviewer 4

My opinion here is partially covered in the response to question 11. The pH of the sediment is only of consequence to the degree that it influences the pH in the water column at the sediment-water interface. I am not aware of any methodologies, other than direct measurement, that would allow estimation of the interface pH if the only data available were the sediment pH and the water-column pH.

5.6.4.1. EPA Response

See response 5.5.1.1.

5.6.5. Peer Reviewer 5

In addition to the aforementioned comments regarding the need for additional studies involving mussel responses before including a consideration for this difference, site-specific variance procedures inclusive of sediment characteristics for a "mussels present" site might be especially applicable in this instance. Should these site criteria account for such differences in testing of site waters with and without sediment, they would be expected to be applicable for the species (or suitable surrogate) and sediment of interest. Use of a site-specific criterion for copper that incorporated such broad test considerations has been utilized to insure protection of mussel species for almost two decades within the Clinch River, Virginia (Van Hassel, 2007).

5.6.5.1. EPA Response

See response 5.5.1.1.

5.6.6. Peer Reviewer 6

Same comment as for Question 11.

5.6.7. Peer Reviewer 7

The freshwater ammonia criteria must appropriately address the pH of the juvenile freshwater mussel environment to be reliable. Although the criteria must be protective they must also be representative of aquatic life sensitivities and implemented in a way that does not overstate those sensitivities. Being protective is straight-forward using conservative assumptions but criteria development and implementation already have multiple layers of conservatism. These layers include the return frequency (once in 3 years) of the criteria, the assumption that exposure is continuous instream for a time period equivalent to the duration of the tests, an acute criterion averaging period of one hour, the use of the 90th percentile of temperature and pH to apply the ammonia criteria, the reasonable potential process of EPA's Technical Support Document for Water Quality-based Toxics Control which relies on a very high percentile (95th or 99th) of the distribution of the highest concentration measured for an effluent to issue permit limits, the 303(d) listing process which is dependent on the vast minority of data exceeding the criteria to conclude that a water body is not in compliance with standards, etc. The conservatism built into criteria development and implementation demands that pH instream be accurately assessed and applied to the criteria to reliably determine compliance with water quality standards.

EPA has not adopted data quality objectives to limit the amount of uncertainty that can exist when applying water quality criteria instream, therefore a consensus standard to determine if water based data can be used to estimate sediment based responses is not available and scientific defensibility cannot be assessed. However, there is an approach that will likely provide consensus agreement in scientific defensibility: to directly address the pH of the instream sediment within the context of a toxicity test rather than trying to compensate for pH differences between water tests and the instream sediment. This approach will directly address concerns that juvenile test results are a function of the presence/absence of sediment and avoid the uncertainty of mechanisms attempting to extrapolate response from water only exposures. It is recognized that this recommendation would not allow the use of current data to derive ammonia water quality criteria inclusive of freshwater mussel sensitivities. It would ensure that the unique ecology of these mussels' life stages is accurately captured (within the constraints of conventional testing methods) and that when the resulting criteria are implemented management actions will be reliable.

5.6.7.1. EPA Response

See response 5.5.1.1.

5.7. Question 13: Should exposure tests on juvenile mussels be conducted with or without sediment in the test chamber?

5.7.1. Peer Reviewer 1

Results of toxicity tests without sediment provide good information regarding the toxicity of ammonia via exposure to ammonia concentrations in water. Acceptable control mortality in a test remains a significant indicator that the test conditions and the general overall health of the tested organisms are reasonable and the increasing mortality observed in the test chambers with higher concentrations of the tested substances is a direct result of the toxicity of the toxic substance. If the mortality in the control chambers remains low then, this is good evidence that the tested organisms are not experiencing unacceptable stress during the test. In several water-only 28-day tests, 100% to 90 % survival was observed in controls. This provides evidence that water-only tests with juvenile mussels can provide acceptable conditions for assessing the toxicity of a contaminant.

Water-only exposure tests provide relatively comparable and reproducible exposure conditions that are useful for assessing toxicity sensitivity and to distinguish effects of pH temperature or hardness etc. Toxicity testing in sediment will not change what we know about the sensitivity of the species to ammonia via a water-only exposure as used and summarized in this criteria document. As shown in Tables 1, 2 and 4, the variability of toxicity values within a species and within genera is all within a factor of 10. Water-only toxicity tests methods have shown good intra and inter laboratory variability for toxicity tests with ammonia, as well as copper and chlorine. This lends confidence in these data for use in criteria development. Toxicity data from water-only exposures are most relevant to the purposes that water quality criteria are most often employed; assessments of natural waters by comparison to the criteria and for use in developing regulatory limits for permitted discharges of the pollutant.

Concerns that tests conducted in sediment might produce different toxicity results are speculation. Some speculation is that water-only exposure stresses the test organisms and may make the test organisms more sensitive to the toxicant. However the good survival in the control chambers indicates that the tested mussels are not overly stressed during the test. Also, test conditions are designed to provide optimum dissolved oxygen and temperature conditions. Natural conditions often are less than optimum with lowered oxygen levels and other conditions. The sediment may be a more natural environment for the juvenile mussels, making them more hearty and resistant to the toxic substance, or the sediment may act to absorb the toxic substance and make it less bioavailable and less toxic. On the other hand, if the sediment may absorb the toxic substance and over time increase the concentration in the interstitial pore

water in relation to the overlying water. Most of these issues are site specific in nature and beyond the scope of national recommended criteria.

Natural sediments are highly variable and even a standardized reconstituted “clean” sediment to be used in laboratory tests may prove variable in some unknown way perhaps in response to some unrecognized factor, however this is simply speculation. Use of sediment in toxicity tests will likely introduce added variability without providing additional value for judging sensitivity and this will introduce added complexity to interpreting the tests’ results for criteria development.

Several studies (Whiteman et. al., 1996, Hickey and Martin, 1999) have reported that ammonia in interstitial pore water can often be higher than concentrations in the water. Overall, these concerns about increased ammonia concentrations in interstitial pore water are not so much related to toxicity sensitivity or issues that directly influence the derivation of criteria. These concerns and observations are more related to an exposure assessment of potential risk to the sediment dwelling organisms and the related physical factors that influence the exposure conditions.

I believe that toxicity data from water-only tests that were conducted under acceptable conditions (with acceptable control mortality) provide the most useful data to judge the quantitative toxic effects of the toxic substance. These water-only data can then be used to assess potential for toxicity in the water column. Currently, there is insufficient information available to indicate that the toxicity of ammonia should be considered fundamentally different in interstitial pore water as compared to the toxicity observed in water-only tests under the same pH and temperature conditions. An ammonia toxicity test with water-only exposures provides reliable test results that can be applied to exposures if the juvenile mussels are exposed to concentrations of ammonia in water. Water exposure tests should be used in the derivation of water quality criteria and this information should apply equally well to water column or interstitial pore water.

If there is a special need to assess interstitial pore water at a site, the criteria can be applied to an assessment of the interstitial pore water. Samples of interstitial water can be collected and analyzed for ammonia concentrations and pH can be measured in the field. If monitoring of interstitial pore water is conducted, care should be taken concerning the depth at which the samples are taken. Juvenile mussels are likely to be within a few centimeters of the surface and juveniles and adult mussels are capable of some control over maintaining their position in the sediment and should not be assumed to be regularly found buried in deeper sediments where low dissolved oxygen, reduced food availability and other stress factors may overshadow risk due to ammonia toxicity. Water samples for ammonia analysis and pH field measurements should be taken at the surface interface or at a depth where mussels are likely to be found.

5.7.1.1. EPA Response

In a recent study, Wang et al. (2011) evaluated the influence of substrate on the sensitivity of two-month-old juvenile fatmucket mussels to ammonia in 28-day water-only exposure and substrate exposure. The water chemistry and pH in the exposures with substrate were measured and relatively consistent before and after passing through the substrate indicating that the substrate was not altering the chemistry. Importantly, the survival response between the water-only and substrate treatments was similar. Dry weight measurements of the mussels increased in the water-only exposure compared to the substrate exposure suggesting that the presence of the substrate increased food availability, as noted by the authors. Based on the apparent improved health of the test organisms in the substrate exposures, and the lack of any significant alteration of water chemistry in the exposure, EPA used the substrate exposure data for calculating the species mean chronic value for fatmucket mussels rather than the water-only exposure data from Wang et al. (2011). See discussion in the 2013 ammonia criteria on p. 35.

The decision to exclude the growth data from the criteria derivation was based on the uncertainty in the test methods for assessing the growth endpoint and the need, as stated by the authors, for additional research “to optimize feeding conditions, to conduct longer-term exposures (e.g., 90 d), and to compare growth effect to potential reproductive effect in partial life-cycle exposure” (Wang et al. 2011). The

growth endpoint showed a high degree of variability, and the test methods for assessing growth, based on substrate or water-only exposures, are currently being evaluated. Wang et al. are conducting further tests on juvenile mussels with substrate over exposures longer than 28-days.

5.7.2. Peer Reviewer 2

If the test is designed to address water column concerns, then the exposure test using juvenile mussels should be conducted without sediment in the test chamber. The use of sediment in water quality toxicity tests introduces too many confounding factors in tests designed to evaluate the needs to be protective of aquatic life in the water column. For many species of mussels, the juveniles primarily feed on particles in the sediment and filter the pore water, rather than directly siphon from the overlying water column. Depending upon the composition of the sediment (i.e., organic matter content) the sediment layer can have profound influences on the pH and ammonia concentration within the pore water of the sediment. In addition, the metabolic byproducts of juvenile organisms that live in sediments also influence the level of ammonia observed within pore waters of sediment. For these reasons alone it is difficult to separate the effects solely due to conditions within the water column versus the effects of the sediment layer.

Furthermore, in static renewal tests that incorporate the use of sediment, there is the potential for decreased ammonia concentrations, as well as pH, in the pore water of sediment when compared to the water column conditions (Allan and Maguire 1992); whereas in natural conditions this relationship is often reversed, such that pH levels and ammonia concentrations are greater in the sediments than the overlying water column. The sediment based toxicity studies of Whiteman et al. (1996) tried to address these gradient issues between the sediment and water column, by introducing nitrogen via sediment and maintaining overlying water concentrations in a non-toxic range. However, the mobility of the test organisms allowed them to seek out the zones of lower ammonia concentration at the interface rather than within the sediment, which also confounded the test results.

This concludes my responses to the Charge Questions. Given the time constraints, I feel somewhat limited in what I could provide. I hope this review was helpful – and thank you for the opportunity.

5.7.2.1. EPA Response

See response 5.7.1.1.

5.7.3. Peer Reviewer 3

Tests using juvenile mussels should be conducted using sediment. Several researchers reported better survival and growth of juvenile mussels with the addition of sediment (e.g., Gatenby et al. 1996, Yeager et al. 1994) and I would note that many of the mussel culturing facilities that produce juveniles for transplants, use sediment in their culturing systems. The ASTM standard for glochidia and juvenile mussel testing (E 24550-06) indicates this in several places including sections 10.4.2.4, 10.5.2.1, 10.5.2.5, 10.6.3.3, and 10.6.3.13. Laboratory studies by Yeager et al. (1994) observed that young juveniles burrow immediately to < 1cm deep in the sediment, and that they were not exposed to overlying water. Others also have reported rapid burrowing of new juveniles in sediment. This information suggests to me that some sediment in test containers may help ensure that juvenile mussels are unstressed and respond appropriately to contaminants. This might be a good check to see whether ammonia concentrations shown to be toxic in water-only exposures are also toxic when the tests are conducted with some sediment (clean, sterile, etc.) as well.

5.7.3.1. EPA Response

See response 5.7.1.1.

5.7.4. Peer Reviewer 4

My comments here are limited to laboratory tests where issues of pH variability between the sediment and the water column could be eliminated (to avoid direct measurement as discussed above) by experimental manipulation. The main reason for including sediment would presumably be to reduce extraneous stress on the experimental organisms, stress which could possibly interact with the toxicant to produce an anomalous result. I have no opinion as to whether the presence or absence of sediment would stress juvenile mussels in an experimental setting. This will have to be determined experimentally.

5.7.4.1. EPA Response

See response 5.7.1.1.

5.7.5. Peer Reviewer 5

If the need is greater at this point with the reassessment to consider data from tests that offer repeatability and precision of results, than to ensure inclusion of test conditions reflecting the life history and habitat selection for this life stage of the mussel, then the test should exclude the sediment considerations. There is a growing body of evidence from tests without sediment that suggests comparable juvenile shell growth can be measured in control treatments in static, static-renewal, and flow-through conditions. This is somewhat contradictory to earlier studies that recognized the importance of sediment in distinguishing effects during longer term exposures with freshwater mussels and fingernail clams. For this reason, the information still seems lacking in determining how much of a confounding effect is introduced by the effects of sediment characteristics (e.g., particle size, microbial mobilization and organic carbon content). More specific information on the survival and physiological condition of benthic organisms when tested in a water-only medium seems necessary for testing beyond 96 hours. Burton et al. (2002) describes a range of toxicity assays using benthic invertebrates in sediment-free systems such as interstitial water, elutriate phase, or spiked waters, with each possibly differentiating toxicant uptake, pathway, and hazard. It would seem prudent to be consistent with current conventional approaches to sediment evaluations for their consideration of appropriate phases to be tested (extractable, elutriate, interstitial, whole sediment, and in situ) relating to appropriate conditions for the species or substrate of interest. The complex interactions between mussel, water, and sediment that influence their responses has been cited to reinforce those considerations (Thorsen et al., 2007).

5.7.5.1. EPA Response

See response 5.7.1.1.

5.7.6. Peer Reviewer 6

I believe this question is very similar to the *Hyalella azteca* question. If adequate test results that are representative to the toxicity of ammonia can be performed in water only, that is how they should be performed. If sediment is necessary for the test to be performed adequately, then they should use sediment. However, it should be cautioned that when the variable of sediment is added to the toxicity testing methodology, the complexity of the test increases considerably as to the interpretation. In my opinion the agency should caution the use of sediment in the test may lead to positive and negative bias of the test results. The agency should try to stay with water column only whenever possible.

5.7.6.1. EPA Response

See response 5.7.1.1.

5.7.7. Peer Reviewer 7

It is clear from Newton and Bartsch (2007) that the potential for juvenile freshwater mussels to be stressed in the absence of sediment exists. Wang et al. (2007) found that sediment did not appear to

impact shell growth and may have slightly reduced survival, but only one test using sediment was conducted and this test only consisted of two replicates per treatment. As stated previously toxicity tests must represent the stress experienced by a single stressor to be used in developing water quality criteria. Therefore for this particular use, until more information is available, juvenile tests of any duration must include sediment in the test vessels in order to be used in deriving water quality criteria.

5.7.7.1. EPA Response

See response 5.7.1.1.

5.8. Additional Comments

5.8.1. Peer Reviewer 3

Other Comments Regarding Appendix A “EPA Final Draft Position Statement on: Acute Toxicity Tests Using Freshwater Mussels”

Under “USEPA’s Aquatic Life Criteria Coordinating Committee” comments, point #2 makes it appear that EPA is willing to alter a national criterion due to a commercially valuable species (e.g., trout) but not an ecologically critical species?

5.8.1.1. EPA Response

This is consistent with Section IV.A (Final Acute Value) in 1985 Guidelines (p. 26), which states: “...if the Species Mean Acute Value of a commercially or recreationally important species is lower than the calculated Final Acute Value, then that Species Mean Acute Value replaces the calculated Final Acute Value in order to provide protection for that important species.”

Note: Appendix A from the draft 2009 criteria is not in the 2013 final ammonia criteria document, rather the current decisions regarding use of acute toxicity data for glochidia and juvenile mussels is addressed directly in the body of the 2013 criteria document.

Point #3: I agree there is no evidence that glochidia in conglutinates are exposed to water-borne pollutants but there’s no evidence that they are not exposed either. Also, uptake of pollutants by planktonic invertebrates lacking gills (e.g., Daphnia) occurs via absorption through the carapace, not just through food.

5.8.1.2. EPA Response

See response 2.3.1.1. The absorption of some pollutants by planktonic invertebrates lacking gills can occur through the carapace.

Point #8: “For glochidia that attach to hosts,” – all glochidia are meant to attach to hosts. What I think you mean here is “In terms of the duration of free-living glochidia, there is a major difference....”

5.8.1.3. EPA Response

Noted, thank you.

Other comments regarding acute data in the Reassessment Document

Daphnia GMAV is Incorrect

I believe the document reports an incorrect GMAV for Daphnia (p. 9) as well as test data in Table 1 (p. 49-50). It appears that the unionized ammonia values reported by Russo et al. (1985) and reported as such in EPA’s 1999 document, were incorrectly transcribed as total ammonia values (see p. 109,

Appendix 4 of 1999 document as well as p. 38, Table 4 of the 1999 document). Even with the temperature correction for invertebrates in the Reassessment Document, *Daphnia* couldn't possibly go from Rank 19 in acute sensitivity to 8th in the draft document. This error has other ramifications as well: assuming the *Daphnia* value is incorrect in the draft, when mussels were absent, the four most sensitive genera are *Corbicula*, *Potamopyrgus*, *Pleurocera*, and *Prosopium* (mountain whitefish). If *Corbicula* is removed along with the unionids as EPA proposed (and which I disagree with), then *Deltistes* (Lost River sucker) becomes the 4th most sensitive genera. Either way, invertebrates no longer occupy the four most sensitive genera. Although I did not do a comprehensive check, it appears that the other acute values in the draft document are consistent with the 1999 document.

5.8.1.4. EPA Response

The acute toxicity values reported for Daphnia magna in Russo et al. (1985) are correct in the 2013 final ammonia criteria document, as is the GMAV for Daphnia. Furthermore, EPA thoroughly cross-checked all values provided in text and in the various data summary tables for consistency in the values provided.

Temperature Equation for Invertebrate and Acute Data should be re-examined

The temperature-ammonia sensitivity relationship applied by EPA to invertebrate acute data in the draft document is based on the slope provided in the 1999 document, based on Arthur et al. (1987), but not used for the acute criterion in that document. As the 1999 document makes clear, the temperature relationships from Arthur et al. (1987) were based on seasonal tests of experimental stream water, including condition, life stage, etc. of test organisms collected during each season. As Arthur et al. (1987) notes, several water quality characteristics other than temperature varied with season and, even more importantly perhaps, organism condition was unlikely to be similar across seasons and at the least was unknown. EPA provided what I believe was a poor rationale for using Arthur's et al. data for chronic invertebrate data. The fact that Arthur did not observe a clear temperature relationship with fish (which are also cold blooded) would weaken the case that temperature was what was driving the apparent ammonia sensitivity in invertebrates in his study (and which Arthur himself acknowledges in his paper).

The temperature-toxicity relationship derived by EPA for ammonia in the 1999 document had a fairly steep slope; e.g., a difference between 20 and 25°C resulted in 2-3 fold decrease in the LC50. There are a few invertebrate species in the draft document for which it appears possible to reanalyze a temperature toxicity relationship, and to determine if one actually exists.

I have not had time to do a formal analysis of the invertebrate acute data, however, my initial reaction is that temperature may be somewhat weakly related to acute ammonia toxicity in invertebrates. For example, for the snail species *Potamopyrgus antipodarum*, Hickey and Vickers (1994) reported an LC50 of 8.71 mg N/L at pH = 8.2 and 15°C and an LC50 of 4.08 mg N/L at pH = 8.2 and 25°C (a factor of 2 over a 10°C difference). The LC50 they reported at the same pH and 20°C was nearly identical to the LC50 at 25°C (4.73 mg N/L). LC50's for the fatmucket mussel at pH = 8.1 and 20°C ranged from 5.2 mg N/L to 11 mg N/L (Wang et al. 2008) a factor of about 2, which could be within the range specified by a 5°C difference in EPA's temperature equation. Even larger differences in LC50's are commonly reported for other invertebrates based on tests at the same pH and temperature (e.g., *Villosa iris* LC50 = 3.0 – 8.9 mg N/L at pH = 8.3 and 20°C). These data suggest that a temperature adjustment factor may be premature for acute invertebrate data. Acute data for *Ceriodaphnia dubia* may perhaps provide the best information for invertebrates thus far. At pH = 8.2 and temperature of 7°C, the LC50 = 16.65 mg N/L (Nimmo et al. 1989). At pH = 8.16 and 22°C, the LC50 = 30.08 mg N/L (Black 2001), however, at pH = 8.0 and 25°C, the LC50 = 14.52 mg N/L (Scheller 1997), similar to the LC50 reported at 7°C. EPA should reanalyze current acute data, similar to the way they did for the 1999 document, to determine the relationship between temperature and acute ammonia toxicity to invertebrates.

5.8.1.5. EPA Response

See response 4.1.2.1.

5.8.2. Peer Reviewer 4

I have reviewed the document listed above as requested in the Charge to Reviewers dated 24 July 2009. In addition I have gone through the background document “1999 Update of Ambient Water Quality Criteria for Ammonia”. Let me state at the outset that I find the Draft Reassessment to be a well conceived, well developed and carefully written document. The acute and chronic ammonia criteria for freshwaters with mussels present, and separate criteria for waters with mussels absent, are well founded and scientifically defensible. I have presented my review comments as responses to the thirteen questions posed in the Charge to Reviewers. Those questions are repeated here for clarity.

5.8.2.1. EPA Response

Noted, thank you.

5.8.3. Peer Reviewer 6

I have found the criteria document to be well organized and very well thought out. I also feel that the scientific justification laid out in the document to be excellent.

I have several questions or comments presented below as well as imbedded in the Charge Questions section (my comments are presented in Bold).

I do question why only standards consistent with ASTM Standards for testing (page 14) are the only relevant test procedures that were used in the document. I believe that the EPA’s Whole Effluent Testing (WET) toxicity testing guidelines for testing are also relevant. Did this exclude any additional data? From my review of the references, I do not believe so.

5.8.3.1. EPA Response

While indication of procedures and other test methods consistent with ASTM or other equivalent “fully-vetted” toxicity test guidance is useful (if not preferred in many cases), the reviewer is correct to say that following such guidance is not mandatory. Instead, per the 1985 Guidelines (Section II, p. 21), the data need only be available with “enough supporting information to indicate that acceptable test procedures were used and that the results are probably reliable...,” as well as meeting the other data quality acceptance criteria provided in the 1985 Guidelines.

I also question when the Nile tilapia was placed in the other chronic toxicity data (and I agree this was appropriate), why is the Asiatic clam included in the data used for development of the criteria (and in fact it is the third most sensitive species for the acute criterion). The Asiatic clam is also a nonindigenous species (or exotic). In my opinion I don’t believe that the Asiatic clam should be used in the development of the criteria regardless that this exotic is found throughout United States. I believe elimination of this species from the dataset will still generate a protective acute criterion.

5.8.3.2. EPA Response

EPA re-considered the inclusion of data from resident non-native or invasive species in criteria development, and concluded, for the ammonia criteria document, that because adequate representation of sensitive taxonomic groups within the overall [acute and chronic] dataset already exists, the additional toxicity data from acceptable tests with non-native or invasive species residing in North America does not improve the criteria dataset(s). For example, EPA removed data for invasive species such as the Asiatic clam from the dataset for ammonia having decided they were not needed as surrogates in this case. Note that the 1985 Guidelines define residency to North America as species which “have reproducing wild populations in North America and have been used in toxicity or bioaccumulation tests... [and that]

Species do not have to be native to be resident.” The intent of considering toxicity and bioaccumulation test results from all species is one of surrogacy with respect to ensuring basic protection for the entire aquatic community.

The Table A in text of the GMAV for the mountain whitefish does not match Table 1. The Text table reports 12.11 and Table 1 reports 12.09. I am expecting that the Table 1 is correct and is what was used for the FAV. Please verify.

5.8.3.3. EPA Response

Noted, thank you. EPA thoroughly cross-checked all values provided in text and in the various data summary tables for consistency in the values provided.