



UNITED STATES ENVIRONMENTAL PROTECTION AGENCY
WASHINGTON, D.C. 20460

OFFICE OF THE ADMINISTRATOR
SCIENCE ADVISORY BOARD

March 8, 2006

EPA-SAB-CON-06-004

The Honorable Stephen L. Johnson
Administrator
U.S. Environmental Protection Agency
1200 Pennsylvania Avenue, N.W.
Washington, D.C. 20460

Subject: EPA Science Advisory Board (SAB) Consultation on a Proposed
Framework for Revising the Guidelines for Deriving Water Quality Criteria
for Protection of Aquatic Life

Dear Administrator Johnson:

EPA's Aquatic Life Criteria Guidelines Committee has developed a proposed framework for revising and updating the Agency's 1985 guidelines for deriving water quality criteria for the protection of aquatic life and aquatic-dependent wildlife ("Guidelines"). The water quality criteria provide guidance to states and tribes for adopting water quality standards. EPA's Office of Water requested that the EPA Science Advisory Board (SAB) conduct a consultation on the proposed framework for revising the Guidelines.

Accordingly, the SAB Aquatic Life Criteria Guidelines Consultative Panel ("Panel") met in a public meeting held in Washington, D.C. on September 21, 2005, to conduct a consultation. The Panel provided comments on: 1) the scope of the proposed framework for revising the Guidelines, and 2) the scientific validity and appropriateness of proposed approaches and methods for developing water-based, tissue-based, and taxon-specific water quality criteria.

In general, the Panel was complimentary of the work undertaken by EPA, and expressed support for the Agency's efforts to incorporate kinetic modeling and a tissue-based approach into the Guidelines. The Panel provided a number of comments and suggestions to strengthen the Agency's proposed framework for revising the Guidelines. The Panel suggested that EPA consider using a tiered approach for completing parts of the proposed revision. The Panel also emphasized the importance of incorporating ecological realism, an exposure component, and field data into the revision process.

There will be no formal report from the SAB since this was a consultation. Individual Panel member comments and recommendations have been included in the minutes of the

meeting and provided to the Office of Water. We hope these comments and recommendations are helpful to EPA as the Agency continues the important work of revising the Guidelines.

Sincerely,

/signed/

Dr. M. Granger Morgan, Chair
EPA Science Advisory Board

/signed/

Dr. Kenneth Dickson, Chair
Aquatic Life Criteria Guidelines
Consultative Panel
EPA Science Advisory Board

**U.S. Environmental Protection Agency
Science Advisory Board
Aquatic Life Criteria Guidelines Advisory Panel**

CHAIR

Dr. Kenneth Dickson, University of North Texas, Denton, TX

MEMBERS

Dr. John P. Connolly, Quantitative Environmental Analysis, LLC, Montvale, NJ

Dr. Frank Gobas, Simon Fraser University, Burnaby, British Columbia

Dr. Christian Grue, University of Washington, Seattle, WA

Dr. Charles Hawkins, Utah State University, Logan, UT

Dr. Michael Hooper, Texas Tech University and TTU Health Sciences Center, Lubbock, TX

Dr. Lynn McCarty, L.S. McCarty Scientific Research & Consulting, Markham, Ontario

Dr. Joseph S. Meyer, University of Wyoming, Laramie, WY

Dr. Judith L. Meyer, University of Georgia, Athens, GA

Dr. Michael C. Newman, College of William & Mary, Gloucester Point, VA

Mr. Robin Reash, American Electric Power, Columbus, OH

Dr. Daniel Schlenk, University of California, Riverside, Riverside, CA

Dr. William Stubblefield, Parametrix, Inc., Albany, OR

Dr. Judith S. Weis, Rutgers University, Newark, NJ

SCIENCE ADVISORY BOARD STAFF

Dr. Thomas Armitage, Designated Federal Officer, U.S. Environmental Protection Agency, Washington, DC

**Summary Minutes of the
U.S. Environmental Protection Agency (EPA)
Science Advisory Board (SAB)
Aquatic Life Criteria Guidelines Consultative Panel Meeting
September 21, 2005 Washington, D.C.**

Panel Members: See Panel Roster – Appendix A

Date and Time: Wednesday, September 21, 2005, 8:30 A.M. – 5:15 P.M.

Location: SAB Conference Center, 1025 F Street, N.W., Suite 3705, Washington, D.C.

Attendees: Chair: Dr. Kenneth Dickson

Panel Members: Dr. John Connolly
Dr. Frank Gobas
Dr. Christian Grue
Dr. Charles Hawkins
Dr. Michael Hooper
Dr. Lynn McCarty
Dr. Joseph Meyer
Dr. Judith Meyer
Dr. Michael Newman
Mr. Robin Reash
Dr. Daniel Schlenk
Dr. William Stubblefield
Dr. Judith Weis

EPA SAB Staff: Thomas Armitage, Designated Federal Officer
Vanessa Vu, Director, EPA Science Advisory Board Staff Office
Anthony Maciorowski, Associate Director, EPA Science Advisory Board Staff Office
Vivian Turner, EPA Science Advisory Board Staff Office

Other EPA Staff: Suzanne Ayvazian
Heidi Bell
Richard Bennett
Walter Berry
Heidi Bethel
Valerie Chan
Luis Cruz
Charles Delos

Other EPA Staff: Russell Erickson
Tim Gleason
Tala Henry
Amie Howell
Lisa Huff
Susan Jackson
Kellie Kubena
Edward Ohanian
Duncan Powell
Mary Reiley
Trish Rider
Cindy Roberts
Keith Sappington
Robert Spehar
Charles Stephan
William Swietlik
Brian Thompson
Glen Thursby
Diane Nacci

Others Participating: Thomas Augspurger, U.S. Fish and Wildlife Service
Janet Burris, Syracuse Research Corporation
Michael Fry, American Bird Conservancy
Chris Hornback, NACWA
George Noguchi, U.S. Fish and Wildlife Service
Richard Schwer, DuPont Corporation
Joe Skorupa, U.S. Fish and Wildlife Service

Meeting Summary

The discussion followed the issues and timing as presented in the meeting agenda (Appendix B)

Convene Meeting, Call Attendance

Dr. Thomas Armitage, Designated Federal Officer (DFO) for the SAB Aquatic Life Criteria Guidelines Consultative Panel opened the meeting at 8:30 a.m. He stated that the Science Advisory Board (SAB) is a chartered federal advisory committee whose meetings are public by law. He reviewed Federal Advisory Committee Act (FACA) requirements, the Panel's compliance with federal ethics and conflict-of-interest laws, and the panel formation process. Dr. Armitage stated that, as DFO he would be present during Panel business and deliberations. He stated that records of Panel discussions are maintained, and that summary minutes of the meeting would be prepared and certified by the Panel Chair. Dr. Armitage then asked the Panel members to identify themselves and their affiliations.

Dr. Vanessa Vu, Director of the EPA Science Advisory Board Staff Office, welcomed the meeting participants and thanked them for providing advice to EPA concerning proposed approaches to revising the aquatic life water quality criteria guidelines.

Purpose of the Meeting and Review of the Agenda

Dr. Kenneth Dickson, Panel Chair thanked the panel members for serving. He expressed his appreciation for the opportunity to provide advice to EPA on development of the aquatic life water quality criteria guidelines. He noted that because the meeting was scheduled for only one day, it would difficult to cover the broad range issues that had been brought before the Panel for discussion. He noted that there would be insufficient time to get into great detail. He stated, however, that the Panel could provide valuable comments and important recommendations to EPA.

Dr. Dickson then reviewed the agenda indicating that the Panel would hear presentations from EPA in four areas: 1) background on and an overview of the proposed revision of the water quality criteria guidelines 2) proposed revisions for deriving water-based criteria, 3) proposed revisions for deriving tissue-based criteria, and 4) proposed revisions for deriving taxon-specific criteria. Following presentations in these areas the Panel would discuss the charge questions on the agenda. He noted that the Panel had been convened to provide a consultation to EPA (not to review a final product) and therefore the Panel would not write an advisory report. Dickson then asked EPA staff for the first presentation on the agenda.

Planned Activities and Overview of Proposed Revision of EPA's Aquatic Life Criteria Guidelines

Dr. Edward Ohanian, Director of the Health and Ecological Criteria Division in EPA's Office of Water provided background on EPA's Aquatic Life Criteria Guidelines (the Guidelines) revision effort. Ohanian's presentation is provided in Appendix D. He reviewed the statutory requirement for developing, publishing, and revising guidelines for the protection of aquatic life. He briefly talked about the existing guidance for deriving aquatic life water quality criteria (the Guidelines) and EPA's plans to revise the Guidelines. Ohanian stated that the SAB had previously concurred that the Guidelines should be updated, and noted that this was a priority for the Office of Water. Ohanian described the committee that is revising the Guidelines and the process that is being followed. He identified the federal partners involved in the process and stated that EPA will bring the Guidelines to the SAB for review when the revisions are complete.

Dr. Tala Henry, of the Health and Ecological Criteria Division in EPA's Office of Water provided an overview of the proposed revisions to the Guidelines. Henry's presentation is included in Appendix D. Henry reviewed prior efforts to incrementally improve EPA's ambient water quality criteria methodology. She specifically discussed the development of: the wildlife criteria methodology used in the Great Lakes Water Quality Initiative, the use of concentration–response modeling in updating the aquatic life water quality criterion for ammonia, the use of population modeling in deriving the aquatic life water

quality criterion for dissolved oxygen, the use of the Biotic Ligand Model in the draft update of the ambient water quality criteria for copper, and the development of tissue-based criterion for selenium. Henry noted that the revised Guidelines will: incorporate the latest scientific approaches into derivation of ambient water quality criteria, be less prescriptive, and provide flexibility for incorporating risk-based approaches, methods, and models. Henry stated that the revised Guidelines would also enhance the ability to make site-specific adjustments. Henry described key issues to be addressed in the revision of the Guidelines and the organization of the Criteria Guidelines Committee. She noted that the Committee was developing guidelines for derivation of three types of criteria: 1) water-based criteria for chemicals for which water concentration is a reasonable predictor of effects, 2) tissue-based criteria for chemicals for which water concentration is not a reasonable predictor of effects (e.g., those that bioaccumulate or for which diet is an important exposure pathway), and 3) taxon-specific criteria needed to provide appropriate levels of protection for specific taxa. Henry described the shared key components of the ecological risk assessment and water quality criteria derivation processes and reviewed the charge questions to the Panel concerning: 1) the use of EPA's Guidelines for Ecological Risk Assessment as an organizing framework for the development of science-based criteria for the protection of aquatic life and aquatic-dependent wildlife, and 2) whether the proposed criteria types and scientific focus for the criteria types are logical and scientifically valid for developing a holistic and integrated criteria framework.

Panel Discussion of the Scope of the Framework for Revising the Aquatic Life Criteria Guidance

The Panel discussed the charge questions that addressed the scope of the Framework for Revising the Aquatic Life Criteria Guidelines (charge questions 1.2 and 1.2 in Appendix E)

A panelist noted that as EPA revises the Guidelines it is important to continue “thinking outside of the box” in order to review and revise water quality criteria using the existing “1985 Guidelines.”

Another panelist stated that EPA should think strategically to develop a roadmap that shows how the Guidelines will be revised in a “planned way.” The panelist noted that the proposed Guidelines revisions were very extensive and that it may take a long time (perhaps 20 years) to accomplish them. The panelist noted that a timeline to accomplish various parts of the revision is needed. Henry responded that EPA is continuing such planning.

A panelist noted that EPA is also working on the development of biological criteria. He stated that, to the extent possible, there should be coordination between activities to revise the Guidelines and activities to develop biological criteria. He stated that the Guidelines revisions and the development of biological criteria may complement each other but could also lead to conflicts. Henry responded that activities to develop both chemical and biological criteria are being conducted by the same branch of the Agency

(in the Office of Water) but these activities have not been closely integrated to date. She also stated that EPA recognizes that such integration is desirable and the Agency is beginning to work toward this end.

A panelist noted that EPA should develop a better “generic” description of why the Agency plans to invest years of work in the revision of the Guidelines. This description should provide an understanding of why the guidelines are being revised (e.g., whether the use of the existing methodology has resulted in cases where there is gross under protection or too much protection)

Discussion of Charge Questions 1.1 and 1.2

The Panel discussed charge questions 1.1. and 1.2. A panelist stated that EPA’s document, *Overview of Proposed Revisions to the Guidelines for Deriving Ambient Water Quality Criteria for the Protection of Aquatic and Aquatic-Dependent Wildlife*, is a readable document. The panelist expressed some concern that the diagrams in the document are based on toxicity testing information, and that there is not enough ecological balance in the proposed revision. The panelist noted that tiered aquatic life use is important in identifying assemblages that can be used as a reference condition and that this concept could be incorporated into the Guidelines revision. The panelist stated that EPA should consider how this concept could complement the toxicity information presented in the document. The panelist also expressed concern about the choice of species to be tested in deriving water quality criteria. The panelist noted that it is important to articulate what the test species represent. Concern was expressed that the organisms tested may not represent the diversity of species exposed to contaminants and that in this regard, some further ecological understanding should be introduced into the document.

Another panelist stated that it is important to ensure that the criteria have an ecological rather than a “laboratory testing” context. The panelist noted that the general approach proposed for criteria development appears to be reasonable, but there is much work to do and it is not clear how and when this work can all be completed. The panelist expressed concern that the document is focusing only on contaminants rather than other important stressors such as habitat degradation.

A panelist noted that it is important to recognize the contribution of the 1985 Guidelines. The individuals involved in developing the 1985 approach did a good job of balancing the available science and the practicality of testing. The panelist noted that the 1985 Guidelines have stood the test of time but that some improvements can be made. He stated that in problem formulation it is important to clearly state the goals of the water quality criteria (e.g., use in NPDES permits as well as other uses). It is important to state that the criteria represent ambient concentration levels below which aquatic life would be safe. The panelist stated the proposed approach to revising the Guidelines should more clearly articulate the uses of criteria values. He also noted that the process to be used for validating the criteria should also be clearly articulated. It is also important to check the existing criteria to determine whether they are “broken.” He noted that EPA has

described an aggressive approach to revising the Guidelines but a tiered strategy is needed to achieve the proposed goals. Some empirical work must be completed to develop the program and in this regard, a public/private effort should continue. The panelist also noted that it is time for EPA to look at developing criteria for some new compounds.

Another panelist stated that the existing guidelines for derivation of the water quality criteria follow the ecological risk assessment process. He questioned how EPA would go about the process of risk management that necessarily focuses on site-specific drivers. He noted that risk management should be a step in the criteria development process. He stated that risk management is the reason for developing water quality criteria and noted that if the risk management goals and objectives were changed, the entire criteria development process would change.

Several other panelists commented on the proposed framework for revising the criteria. One noted that the proposed framework and criteria types are appropriate. However it is important that the underlying effects assessment consider sublethal effects. Another panelist stated that the framework stresses effects but not exposure. He noted that there did not seem to be much emphasis in the draft document on how exposure will be handled in the implementation of criteria. The Panel discussed Figure 1 in the EPA document, *Overview of Proposed Revisions to the Guidelines for Deriving Ambient Water Quality Criteria for the Protection of Aquatic and Aquatic-Dependent Wildlife*. A panelist noted that there are both dotted and solid lines in Figure 1 and commented that it was not clear what the lines mean. He also commented that the text describing the translation between the ecological risk paradigm and aquatic life criteria paradigm was not clear. Another panelist stated that the framework should address mixtures and also include a discussion of biological criteria. A panelist noted that it is not clear what has happened to the idea of sediment criteria, and indicated that this should also be discussed in the framework. Dr. Lynn McCarty noted that in his initial written comments (provided in Appendix E) he included a redraft of Figure 2 in EPA's overview document. He stated that his redraft addresses sediment as an exposure medium. The Panel chair commented that Figure 2 in EPA's overview document is an important diagram and should clearly represent the processes associated with the different types of criteria. He commented that the current Figure 2 in the EPA document "does not do justice to the exposure side" of criteria development. EPA responded that Figure 2 is a communication tool to illustrate how the criteria types relate to each other.

The Chair then stated that the Panel would move to the next topic on the agenda, proposed derivation of water-based criteria, and would return to the discussion of charge question 1.2 if there were additional time available before lunch.

Proposed Revisions for Deriving Water-Based Criteria

Mr. Charles Delos of the Health and Ecological Criteria Division in EPA's Office of Water presented the Agency's proposed approach for deriving water-based criteria (provided in Appendix D). Delos stated that EPA is concerned that the existing approach

for deriving water quality criteria does not consider the effects of duration of exposure in eliciting a toxic effect. Delos described modeling approaches that EPA is proposing to use to improve the derivation of water quality criteria: 1) a kinetic toxicity modeling approach to translate from lab test exposures to continuously variable concentrations, and 2) a life stage structured population modeling approach to account for population reduction from effects on survival and reproduction, and rate of recovery after population loss. Delos provided a detailed overview of the models being considered for use in future water quality criteria guidelines procedures. These included: 1) a first-order toxicokinetic model in which toxicant accumulation is the integral of an uptake rate proportional to the exposure concentration, and an elimination rate proportional to the accumulation; 2) a deterministic process model that would be used to evaluate the influence of accumulation on organism survival; and 3) a stochastic process model that would be used to evaluate the influence of accumulation on organism survival.

Delos also stated that EPA wanted to define water quality criteria exceedance in terms of the number of taxa affected and how long it takes to replace lost individuals. The Agency has therefore decided to pursue the use of stage-structured population modeling approach in developing water quality criteria. Delos described two population modeling approaches that the Agency is considering: a density independent approach, and a density dependent approach. Delos then reviewed the charge questions to the Panel.

Discussion of Charge questions 2.1 – 2.3

The Panel discussed charge questions 2.1 – 2.3 (provided in Appendix E). These questions focused on whether the EPA's proposed kinetic toxicity models, population models, and approach for aggregating effects across species were scientifically appropriate for deriving water-based criteria.

A panel member noted that a bioconcentration factor is a term used in the toxicokinetic model. He asked EPA whether this is a steady state term. EPA staff responded that it is a steady state term and noted that the Agency wants to relate the accumulation of a chemical to a weighted average water concentration. Use of the weighted average concentration is a way to integrate over time.

A panel member asked EPA staff how transferable toxicity tests endpoints using growth were for use in the kinetic toxicity model. EPA staff responded that the data were very transferable. EPA staff stated that the Agency has been looking at multiple toxicity tests with the intention of using these data. EPA staff stated that survival and mortality data were presented to the Panel as examples of information that could be used in the model, but data representing other endpoints could also be used.

Another panel member asked how EPA would decide which effect is sufficiently adverse to translate into water quality criteria concentrations. EPA staff responded that in making such determinations the Agency wants to consider protecting species assemblages but that they would retain the ability to look in detail at the severity of an effect. EPA staff noted that when the existing criteria guidance document was developed a decision was

made to protect 95% of the taxa represented by species tested. Similar decisions would be made using a new approach.

A panel member asked EPA staff whether each species is considered to be independent in the proposed approach or whether there would be any consideration of interaction among species. EPA staff responded that species are considered independently.

The Panel discussed the modeling approaches proposed by EPA. A panel member stated that, in general, she applauded EPA's attempt to incorporate kinetic modeling into the derivation of water quality criteria. However she posed several general questions about the approach. She noted that metals concentrations are correlated with the hydrograph and questioned how such a relationship could be incorporated into the proposed approach. She noted that some populations exist well below carrying capacity and stated that in these populations the models might be over or under protective. She also noted that it would be very important to use unbiased datasets in deriving the criteria. If data from tests with relatively insensitive species were used the criteria would be under protective.

Dr. Michael Newman, a panelist, stated that he thought EPA's proposal to use kinetic toxicity modeling was a step in the right direction. He offered a number of comments on the proposed approach. He stated that:

- There is an enormous literature on this approach in the medical sciences. Models have been in use for a long time in the field of pharmacology. He stated that there are many toxicodynamic models that could be used and much of this literature has not been integrated into EPA's proposed approach. The statement in EPA's documents that toxicodynamic models are not useful is not accurate.
- He noted that toxicodynamic models could be linked to demographic models.
- He noted that in moving to this approach from the use of LC50s there are likely to be problems linking new information with what has been done in the past.
- He stated that there are cases where concentration is not relevant (e.g., where there is oxidative damage to the gills the concentration in fish is not relevant) and this would not be taken into consideration using the proposed approach.

The panelist noted that he had expressed other concerns in his written comments (included in Appendix E).

Another panelist stated that EPA's proposed approach represented a vast improvement of the existing water quality criteria guidance. He noted that many criteria are now based on speculative assumptions. The existing methodology works well for fast acting toxicants, but kinetic models would address the mode of action for other groups of pollutants. The panelist provided several comments on the models.

- Kinetic models assume intermittent exposure, but some waterbodies have constant exposure to pollutants and others have episodic exposure. He noted that there are ranges of exposure scenarios that EPA should model.

- He noted that he did not see problems with the use of deterministic population models. However, the biggest problem will be parameterization. He noted that many populations have high variability in population sizes.
- He noted that in applying the models it is important ensure “ecological reality” by looking at community function to determine the sensitivity of the community to loss of species. The toxicity-effects database is a surrogate for this.

Another panelist expressed strong support for the use of the kinetic toxicity model. He noted the importance of: cross-species extrapolation, the application of structure-activity relationships, and examination of acute vs. chronic effects. He stated that he was not sure how non-toxicant stressors should be incorporated into the approach.

A panelist stated that the proposed use of population models is a sound and forward-looking approach but recommended that that EPA expand the approach to look at metapopulation dynamics. He noted that EPA had proposed using a stage-based approach to population modeling and stated that an age-specific approach might also be considered. He further stated that EPA had not included any discussion of sensitivity or elasticity analysis in the proposed modeling approach.

Another panelist provided additional comments on the modeling approaches. He stated that the water-based criteria approach would be applied to many chemicals. He recommended that EPA use a modeling approach that would enable the Agency to look at net exposure in the field (i.e., use bioaccumulation factors [BAF] rather than bioconcentration factors [BCF]). He noted that bioavailability of contaminants is an important issue that should be considered in any approach and he pointed out differences between BAFs and BCFs. He also identified a number of issues and uncertainties that should be considered when using population models. He noted that stressors vary from site to site. He stated that different stressors (other than chemical stressors) act on different populations and they must also be considered. This can make it difficult to develop one number.

The Panel chair stated that is important for EPA to consider how the Agency would deal with uncertainties in setting thresholds and making decisions.

A panelist stated that EPA should try to develop criteria that protect an assemblage. He asked EPA staff whether they had considered using an approach that directly considers measures of assemblage rather than trying to model the populations comprising the assemblage. In this regard, he suggested that the development of “field-based” criteria might be considered. EPA staff responded that this is an interesting idea but because there are so many confounding factors it is difficult to identify and link causes and effects. The panelist noted that there are strong compensatory responses in invertebrate communities and streams. He noted that EPA needs to reconcile the results of a model with what is seen in the real world.

Another panelist stated that in the real world there are species interactions that should be considered. For example, there may be an increase in the population size of a prey

species if the predator species is affected by a particular stressor. She noted that EPA should consider ecological predator/prey interaction. EPA staff asked the Panel members whether they were recommending a model that could be used. Panel members responded that they were recommending that EPA take these factors into consideration, be creative, and put ecological reality into criteria development. They stated that population interactions should be considered. The Chair of the Panel noted that the proposed models are useful tools but field data results might also be considered. EPA staff responded that the criteria are considered to be indices of risk.

Another panelist stated that when using population models, EPA should find ways to take different age structures into consideration. The panelist also stated that it would be important to consider delayed effects of chemicals like dioxin.

Another panelist expressed concern about using different endpoints in the models. He noted that reproduction, survival, and growth are all toxicity test endpoints. He stated that it is important to understand the toxicity test data used in the models. EPA staff responded that the toxicity tests have limitations and that the data need to be validated against the model.

A panelist observed that an objective of revising the water quality criteria guidelines is to increase accuracy and reduce the uncertainty of the criteria. He expressed concern that uncertainty is introduced when population models are used. He questioned whether the use of such models would actually reduce uncertainty. He stated that there should be an assessment of this, and that the assessment should be documented.

Another panelist stated that in developing the approach for revising the Guidelines, EPA should consider the literature on community interactions. He noted that there are classic ways of looking at communities to determine their stability. He suggested that the use of mesocosms is an approach that would allow EPA to look at community stability and evaluate species-community relationships. EPA staff responded that the Agency is revising the guidelines to improve the current approach for deriving water quality criteria. The Agency has not included species interactions in the proposed approach, but as the guidelines are revised EPA wants to promote the development of new information. EPA staff stated that the Agency would like to lay the foundation for future work not being done at the present time. A panelist responded that EPA should establish connections with other disciplines and proactively determine where the Agency wants to be in the future, not just undertake work that can be accomplished with expected resources. The proposed approach to revising the guidelines should define possible work that might be completed if additional resources were available as well as work has been planned with expected resources. Another panelist suggested that EPA might consider undertaking work to conduct lab to field extrapolation.

Discussion of Charge Question 2.4

The Panel discussed the response to charge question 2.4 (focusing on whether the framework being considered by EPA for deriving water-based criteria is scientifically appropriate for use in deriving the criteria).

A panelist commented that the proposed revision of the water-based criteria is a daunting task. She stated that it is important to incorporate ecological knowledge into the process but acknowledged that this will be difficult. She noted that EPA should think about whether the right questions are being asked. EPA should consider the real problems and major uncertainties to be addressed. She mentioned the following uncertainties and issues that should be considered:

- Water quality criteria uncertainty associated with nonequilibrium conditions. There is at least a factor of 10 difference in water quality criteria uncertainty due to nonequilibrium conditions.
- Addressing multiple chemicals and mixtures. There is a large amount of uncertainty associated with this issue.
- Need for a large amount of data. She noted that the proposed approach to revising the Guidelines was very “data-intense” and it is not clear that EPA will be able to meet its proposed goals in the near term. She stated that EPA might want to define a fallback position if the proposed ambitious goals cannot be accomplished. She questioned whether the proposed revision would accomplish more than addressing the questions of chemical equilibrium and mixtures.

Another panelist stated that the aquatic life criteria are one of several tools used by the Agency to implement water quality protection programs. He noted that there are several other tools to address mixtures. These tools include biocriteria and whole effluent toxicity testing. He noted that the Panel should not ask EPA to revise the water quality criteria to correct problems that other tools may address.

A panelist stated that EPA could not address all of the needed improvements in the criteria at once. He stated that EPA should first look at the use of toxicodynamic models and then move on to toxicokinetic models.

Another panelist stated that it is scientifically appropriate to use a kinetic toxicity modeling approach in the Guidelines revision. He stated that he was not sure whether it is scientifically appropriate to do population modeling. It is important to look at the population models that are available and ask how good they are. It is necessary to understand how much uncertainty is associated with the models before it is possible to judge whether it is scientifically defensible to use the models.

A panelist commented that EPA should be encouraged to develop long and short-term plans for improving the water quality criteria. EPA should build on the 1985 criteria guidance. Pharmacokinetic modeling can be accomplished in the near term; population modeling may be a longer-term effort. The panelist suggested that EPA might want to conduct a hind cast validation to determine whether the current water quality criteria

approach is “broken” and needs improvement or whether more effort should be focused on other projects.

A panelist questioned the meaning of “broken” in the context of the water quality criteria. She noted more attention should be focused on considering chronic sublethal effects. Another panelist noted that multiple exposure pathways are not considered in the 1985 criteria methodology so it is important to think about dietary exposure and developing tissue-based criteria. The Panel Chair expressed the opinion that the water quality criteria methodology is not broken but EPA has undertaken an evolving process to improve the criteria. He noted that a tiered approach to revising the Guidelines should be considered. The Chair then recessed the meeting for lunch and stated that the Panel would reconvene at 1:00 p.m. to hear public comments and continue to discuss the next topic on the agenda.

Public Comments

The Chair reconvened the Panel at 1:30 p.m. and called for public comments. Richard Schwer of DuPont Corporation provided comments. Mr. Schwer stated that he appreciated the opportunity to comment on EPA’s proposed framework for revising the Water Quality Criteria Guidelines. He stated that after a period of 20 years it was timely and appropriate to make sure that the Water Quality Criteria Guidelines represent the state of the science. He stated that he was pleased to see higher visibility given to exposure as well as effects of toxics. He stated that it is important to develop a robust approach to considering exposure, and noted that exposure factors differ from waterbody to waterbody. He also stated that as the Guidelines are revised it is important to keep in mind the fact that states must promulgate criteria values in their water quality standards, and also to remember that the criteria are used for many purposes. He expressed the opinion that it would be important to move gradually to improve an approach to developing water quality criteria that has been successful in the past. He expressed concern that unless more data were provided it would not be possible to develop robust criteria under EPA’s proposed approach. He reminded the Panel that the SAB had reviewed the wildlife criteria developed under the Great Lakes Initiative (GLI) and suggested that the Panel look at the findings of that review. Schwer also expressed concern about the proposed taxon-specific method of deriving water quality criteria. He stated that this method could result in more stringent criteria that may not be necessary, and commented that the method should only be used where appropriate. He also noted that EPA has been involved in an effort to ensure that water quality criteria meet the requirements of the Endangered Species Act, and questioned how that effort fits into the taxon-specific criteria proposal.

The Chair thanked Mr. Schwer for his comments and stated that the Panel would spend some additional time discussing the response to charge question 1.2 (focusing on whether the proposed criteria types are logical and scientifically valid) before moving to the next topic on the agenda, the proposed approach for deriving tissue-based criteria.

Continued Discussion of Charge Question 1.2

A panel member commented that she was very pleased to see that tissue-based criteria were included in the proposed approach. The member asked EPA staff why sediment criteria were no longer part of the proposed approach. EPA staff responded that in the conceptual model sediment is considered as an exposure medium, and Biota Sediment Accumulation Factors (BSAFs) could be used to translate between tissue concentrations and sediment concentrations.

A panelist suggested that in the criteria planning documents, EPA should to discuss how the criteria would be implemented across programs. For example, EPA should discuss how the criteria would be applied in developing Total Maximum Daily Loads and in water quality permitting.

Another panelist stated that EPA's proposed approach did not appear to address mixtures. The panelist recommended looking at mode of action rather than toxicity.

A panelist stated that he did not understand why taxon-specific criteria should be differentiated from water-based criteria if EPA's objective was to ensure the health of aquatic systems. There was agreement on this point from another panel member who stated that the development of the different criteria approaches might confuse people. He noted that in concept, EPA appears to be trying to bring everything together. He stated that "parceling out" the criteria into different types might not work well. Another panelist stated that the water column and tissue criteria did not appear to capture exposure routes of toxics and that this was a significant issue. The panelist stated that development of tissue-based criteria is scientifically valid and logical but it is not necessary to "force fit" tissue-based criteria to chemicals that do not bioaccumulate. He noted that there are many different opinions on what is bioaccumulative and EPA may "put itself into a corner" by making such a distinction.

The chair thanked the panel members and called for the next presentation on the agenda, the derivation of tissue-based water quality criteria.

Proposed Derivation of Tissue-Based Criteria

Mr. Keith Sappington and Dr. Richard Bennett of EPA's Office of Research and Development presented an overview of EPA's proposed approach for deriving tissue-based water quality criteria (presentation is included in Appendix D).

Sappington stated that EPA's existing guidelines do not comprehensively address ecological risks from bioaccumulative chemicals and therefore a tissue-based approach has been proposed by EPA. The rationale for a tissue-based approach and current challenges associated with a tissue-based approach were described. The primary components of the tissue-based approach were identified. These components include procedures for deriving a national tissue criterion and procedures for translating a national tissue criterion into concentrations in media and components of the aquatic food web.

The process of deriving a tissue criterion for aquatic life was described and a range of issues to consider in problem formulation, the analysis plan, and characterization of effects (i.e., setting the criterion) were discussed. It was stated that the criteria might be deterministic (based on appropriately sensitive species) or probabilistic (based on a specified percentile). Issues associated with each of these approaches were identified.

Sappington discussed the translation of tissue criteria into media and food web concentrations. He stated that EPA would propose procedures for translating the criteria into environmental concentrations. He noted that a number of issues must be addressed in the translation process. Sappington stated that in the translation process EPA was proposing to select representative species and use methods for estimating species-specific bioaccumulation potential. Representative species could be defined for a range of exposure potentials within an assemblage or defined on a site-specific basis. Sappington stated that EPA was considering using a bioaccumulation framework similar to that in EPA's human health criteria methodology. He noted that site-specific estimates of bioaccumulation would be encouraged and that nationally representative parameter values might be developed for use when site-specific data are lacking. The output of the process would be multiple translated criteria (i.e. criteria concentrations in water, sediment, algae/macrophytes, zooplankton, macroinvertebrates, and forage fish) for each representative species.

Richard Bennett described a proposed framework for developing tissue-based wildlife criteria. He noted that the process was conceptually similar to the approach for developing tissue-based aquatic life criteria. However, there were differences in the toxicity data available, exposure pathways of concern, and life history of organisms. Bennett noted that where data were available, probabilistic methods could be used to derive tissue-based wildlife criteria, when data were limited deterministic methods would be more appropriate. He noted that national level tissue-based wildlife criteria would be derived and that the national level criteria might be modified at state or local scales if sufficient additional information were available to improve the characterization of risk. The Great Lakes Water Quality Initiative approach for deriving wildlife criteria values was described as a method to be built upon. A summary of issues concerning the development of wildlife criteria was provided. Bennett stated that: wildlife criteria may be based on chemical concentrations in wildlife tissues or diet, the criteria may be calculated using deterministic or probabilistic methods, EPA is focused on a national level approach that could be refined at smaller scales, and methods are being developed to translate wildlife diet or tissue values into concentrations in the aquatic food web.

Discussion of Charge Questions 3.1, 3.2, and 3.3

The Panel discussed charge questions 3.1, 3.2, and 3.3 (focusing on the rationale and conceptual approach for development of tissue-based criteria, the strengths and limitations of the flexible approach used to derive tissue-based criteria, the rationale used by EPA for determining when to use population modeling in the development of tissue-based criteria.)

A panelist expressed the opinion that the proposal to derive tissue-based criteria is the best part of EPA's conceptual approach to revising the Guidelines. He expressed strong support for the use of tissue based-criteria and noted that the approach should not be limited to bioaccumulative chemicals. He stated that it is important to consider dietary concentrations of contaminants. He noted that the proposed definition of bioaccumulative contaminants (those with a log K_{ow} of > 5) is not the right cutoff value. He noted that it should be a much lower value. He stated that this is because there are chemicals with lower K_{ows} that do not bioaccumulate in "water breathing" organisms, but are bioaccumulative in "air breathing organisms." The panelist commented that he had reservations about the application of the species sensitivity distribution. He noted that a selected percentile that is statistically defensible might be ecologically unjustifiable. He noted that use of uncertainty factors might be considered to address this concern. The panelist commented on the extrapolation of contaminant concentrations between tissues. He noted that for many hydrophobic organic chemicals, a lipid-normalized concentration is a better alternative than whole body concentrations in the derivation of tissue-concentration response relationships.

Another panelist commented that the proposed approach to deriving tissue-based criteria had been very well "thought out." He noted that metals might pose a problem in some areas and that an approach to dealing with metals should be considered. He stated that the proposed K_{ow} cutoff for defining bioaccumulative contaminants is "out of the ballpark" for waterfowl. He also noted that one of the greatest obstacles to overcome is lack of data. He expressed the opinion that some regulatory mechanism should be provided to develop data. Additional data must be generated from standard test procedures. He recommended that EPA develop a position paper discussing data needs. The panelist also stated that considering the mechanism of action is important but there are some contaminants where a good correlative relationship between tissue levels and effects can be developed without knowledge of the mechanism of action (he mentioned lead as an example of this). The panelist also stressed the importance of looking at levels of the metabolites of contaminants.

Another panelist provided additional comments. He stated that:

- It is not clear to the general public how the proposed criteria would be used and additional guidance is needed.
- It is surprising that a greater emphasis has not been placed on "non-traditional" endpoints, specifically "sublethal" effects such chemical-induced changes in behavior which can decrease survival, interfere with reproduction, and lead to changes in community structure at concentrations much lower than those associated with direct mortality of adults or embryos.
- Without a good understanding of the factors governing exposure in the field, it will be difficult to predict what an individual's actual exposure will be.
- The toxicity associated with chemical mixtures also continues to be under emphasized or ignored in the proposed criteria approach.
- The availability of appropriate data continues to be problematic. He questioned how the Agency was planning to address these data gaps. He stated that resources

must be made available to obtain the appropriate data. Good time series and behavioral information are needed.

- There is a potential need for more local criteria, or at least guidance from which more local criteria can be developed.
- The panelist agreed with the rationale for developing tissue based criteria, the application of the risk assessment framework, and expressed support for the harmonization of proposed tissue based criteria guidance with bioaccumulation assessment guidance for human health.
- The panelist noted that that sacrificing aquatic-dependent wildlife to obtain tissue residues may be problematic.
- The panelist suggested that EPA use the available data to generate specific case examples of the proposed conceptual approach. He noted that this would be a very important next step.
- The panelist noted that population models can be applied if the information is available, but an apparent lack of life stage specific data appears to be problematic.

Another panel member provided additional comments. He stated that:

- A considerable amount of residue data can be obtained from toxicity tests if the tests were conducted under steady state conditions. The bioconcentration factor (BCF) can relate water exposure to whole body exposure. If BCF data are available one can derive tissue levels by analyzing data sets that have already been collected.
- The differences in species sensitivity can often be related to differences in test protocols.

Another panel member offered additional comments. He stated that:

- The arguments provided by EPA to support use of tissue-based criteria apply to a much broader range of chemicals than those for which multiple routes of exposure are important. He stated that there is a need to better articulate the reason for restricting the use of such criteria to chemicals “with a high propensity to bioaccumulate.”
- Data needs for developing tissue-based criteria should be more clearly identified in order to determine whether it is practical to proceed with the development of the criteria. It is important to consider the level of effort and time necessary to generate the required data. EPA should answer the question, “are tissue-based criteria practical now, in five years, or even in ten years?”
- The translation of tissue criteria to concentrations in water and the food web is flexible given the highly site-specific nature of bioaccumulation and the absence of site-specific data for some chemicals and many sites. The panelist agreed in general with the stated preference to use site-specific data where available and food web modeling where such data are not available. The panelist stated that EPA should develop comprehensive procedures for the compilation and use of data and for the development of a food web model(s).
- Uncertainty estimates should not be used to drive a profoundly conservative criterion. The panelist noted that in this regard, most states do not allow a mixing

zone for bioaccumulative chemicals. The panelist stated that uncertainty estimates should be used to judge whether a tissue-based criterion is inferior to a water-based criterion.

- In the context of population modeling, there appears to be little residue-response information available for integrating responses of various demographic parameters over multiple life stages, such as fecundity and adult, juvenile, and larval survival. Most of the available data are for “legacy pollutants.” Consequently, it is not clear whether it would be feasible or useful to integrate population modeling into national-level tissue criteria for bioaccumulative chemicals.

The Panel Chair stated that ecological realism should be part of the whole criteria package. It is important to provide guidance to States and others concerning the implementation of the criteria.

A number of other comments were provided on the tissue-based criteria approach.

A panelist commented that the proposed approach should consider sublethal endpoints. The panelist stated that such data are available for legacy chemicals. Information on endocrine disruptors and immune system effects is available in the published literature. The panelist stated that this information should be used in EPA’s framework for revising the criteria. EPA staff responded that to be useful in developing tissue-based criteria, the data must tie tissue levels to contaminant effects. Another panelist stated that additional information is available from studies of low levels of exposure (below levels that have previously been considered to be effects concentrations). EPA staff again noted that to be useful, measured tissue data must be associated with an effect.

Another panelist commented that it is important to understand the mechanisms of action of contaminants. He noted that there is wealth of fate, transport, and biological information that can be useful in this regard. EPA staff commented that it is often difficult to look at how effects work through mode of action. The panelist suggested that EPA look at effects on target organs. He stated that it is critical to understand “where the compounds go.” The panelist also commented that he was not convinced that the population model should be used in the tissue-based criterion approach.

A number of panelists discussed how EPA might obtain tissue data without killing target organisms. It was suggested that techniques such as measuring contaminants in bird feathers might be useful.

The Chair thanked the panelists for their comments and called for discussion of the next topic of discussion on the agenda, the approach for deriving taxon-specific criteria.

Discussion of Taxon-Specific Criteria

Mr. Brian Thompson of EPA Region 5 and Mr. Tom Augspurger of the U.S. Fish and Wildlife Service presented an overview of EPA’s proposed approach for deriving taxon-

specific water quality criteria. Thompson stated that risk managers want to protect “special status” species that are known to be sensitive to pollutants and potentially under-protected by the national criteria. Special status species include federally listed threatened and endangered species. He stated that the taxon-specific approach is still being developed and is now in the “conceptual phase.” He provided a species sensitivity distribution to illustrate the statement that a risk management target other than that used in a national criterion may be desired for certain special status species.

Thompson stated that the format of the taxon-specific criteria would not be certain until the format of the national aquatic life criteria is established. The taxon-specific criteria could be actual numbers or a process for deriving numbers. He noted that the issue of multiple stressors would have to be folded into the process.

Thompson and Augspurger presented issues to be considered in developing an approach for deriving taxon-specific criteria. These included: toxicological data quality and quantity (including use of data for surrogate taxa, acceptable data points, and minimum data requirements), ability to populate models, data preferences for surrogate taxa, the goal for level of protection, and consideration of a taxon’s ability to tolerate risk. Thompson presented five approaches for deriving taxon-specific criteria. Thompson stated that when toxicological data exist for a sensitive species and chemical, the method would most likely be based on the same methods used for deriving aquatic life or aquatic-dependent wildlife criteria. When data for the sensitive species are not available surrogate taxa information would be used. The methods for using surrogate data include:

- Use of regression models to develop interspecies correlation estimates
- Use of species sensitivity distributions (deriving a probability density function using available data within a given taxon)
- Looking at variability within a taxonomic level (determining the average variability at a given taxonomic level)
- Use of surrogate data from the most sensitive species within a group

Thompson and Augspurger also discussed the use of uncertainty factors to compensate for lack of knowledge on contaminant sensitivity of various taxa.

Panel members provided a number of comments on the proposed approach. A panel member stated that EPA should consider effects on different life stages. The panel member noted that it is important to consider delayed effects from one stage to another.

A panel member asked where exposure assessment is included in the development of taxon-specific criteria. EPA responded that exposure assessment is addressed in the “toxicity aspect” of the criteria. EPA will look at species sensitivity distributions and let the biology inform the decision.

The Panel provided a number of comments on the use of species sensitivity distributions and interspecies correlation methods.

A panelist asked how species sensitivity distributions would be applied and whether data from the same genera or families of organisms would be used. EPA staff responded that this would depend upon the species of concern. If data were available for a genus within the same family as the species of concern these data would be used. Data preferences would be identified. The first preference would be to use data within the same genus. If these data were not available it would be necessary to move to a higher taxonomic level. The panelist stated that EPA should develop a “preference tree” describing the data to be used.

A panelist noted that unrelated species were used in EPA’s examples of interspecies correlation estimates. The panelist asked whether better estimates could be obtained using data from related species. EPA staff responded it might be necessary to use data from unrelated species because these are often the only data available.

A panelist noted that in EPA’s example of interspecies correlation the relationship was linear. She asked EPA staff whether the Agency is certain that all of these relationships are linear. EPA staff responded that there might be some relationships that are not well correlated. The panelist noted that unless a good correlation can be shown using a large amount of data, the relationship might not be valid.

Another panelist expressed concern about the use of species sensitivity distributions for deriving taxon-specific criteria. He noted that there is quite a bit of variability in the toxicity test data (the data often vary by a factor of 5-10). He noted that that the range of natural variability may make it difficult to derive the taxon-specific criterion. A panelist noted that application of the species sensitivity distribution is different in the three proposed approaches for deriving water quality criteria. This might be difficult to explain to regulatory agencies. EPA staff responded that in the water-based approach and the tissue-based approach the Agency is trying to integrate available information, but in the taxon-specific approach it is necessary to look at single species.

Another panelist noted that EPA had asked the Panel to comment on the scientific defensibility of taxon-specific approaches. The panelist asked EPA staff whether there were issues of particular concern in this regard. EPA staff responded that one issue of concern is variability. The Agency was interested in the Panel’s views on which approach might offer less variability.

A panelist responded that it is important to determine whether the endangered species data show that chemical contamination is contributing to the decline of the species. For salmonids, contamination is less risk than other factors like habitat modification. U.S. Fish and Wildlife staff responded that freshwater mussels and other endangered species are sensitive to metals and ammonia. For most endangered species the problems are multiple stressors.

Another panelist noted that it is important to distinguish between sensitivities to stressors and indications in the field that pollutants are causing low population sizes of the species.

In the problem formulation stage of criteria development it is important to justify that chemical stressors are potentially important factors.

A panelist commented that use of the population model would appear to be an important part of the taxon-specific criteria development process. Another panelist noted that a relative change in abundance in a density dependent model should be interpreted differently for abundant species and endangered species. The change of concern should be much lower for endangered species.

The Panel discussed a number of implementation issues. A panelist asked how the taxon-specific criteria would be implemented in water quality standards. EPA staff responded that it would be important to develop a simple methodology that states can use to develop the numbers. A panel member asked whether the development of taxon-specific criteria would make the criteria for protection of species assemblages moot. EPA staff responded that the taxon-specific number would only affect sites where there are species of high sensitivity. Another panelist noted that some species play a special role in the ecosystem and questioned whether taxon-specific procedures would be applied for protection of these species. EPA responded that ecologically important species are not designated as special status species for which the Agency would derive specific criteria, but states can choose to use taxon-specific methods for any species.

The Chair stated that due to time constraints he would have to end the discussion, thanked the panel and speakers for their contributions, and provided a summary of the discussion.

Summary

The Chair stated that the Panel had provided good advice to EPA on proposed approaches for revising the water quality criteria guidelines. He noted that the process is evolving, that he appreciated EPA's efforts, and that he did not want the Agency to think the Panel was trying to be too critical. He stated that this effort was very important because it would result in improved methods to protect natural resources. He identified a number of general themes that EPA should consider as the guidelines are revised:

- Support was expressed for the use of EPA's Ecological Risk Assessment Framework/Paradigm.
- EPA should develop timelines for completing various parts of the proposed revision.
- The Agency should consider using a tiered approach for implementing the revisions (e.g., consideration of using a pharmacodynamic approach in a first tier)
- As the Guidelines are revised the Agency should consider how the revisions could decrease uncertainty.
- The Agency should consider how ecological realism could be incorporated into all parts of the Guidelines revision process. It is important to learn from available field data.
- The problem formulation step is the most important part of the process.

- The exposure component should be addressed in the process. EPA should consider how environmental factors mediate effects.
- Support was expressed for the development of tissue-based criteria, but EPA should better harmonize the tissue- and water-based approaches
- Case studies should be developed.

The Chair then adjourned the meeting.

Respectfully Submitted:

/Signed/

Thomas M. Armitage, Ph.D.
Designated Federal Officer

Certified as True:

/Signed/

Kenneth Dickson, Ph.D.
Panel Chair

APPENDICES

Appendix A: Roster of SAB Aquatic Life Criteria Guidelines Consultative Panel

Appendix B: Meeting Agenda

Appendix C: Initial Panel Responses to Charge Questions

Appendix D: EPA Presentations

Appendix E: Charge Questions to the Panel

Appendix A – Panel Roster

U.S. Environmental Protection Agency Science Advisory Board Aquatic Life Criteria Guidelines Consultative Panel

CHAIR

Dr. Kenneth Dickson, Professor, Institute of Applied Sciences, University of North Texas, Denton, TX

MEMBERS

Dr. John P. Connolly, President/Senior Managing Engineer, Quantitative Environmental Analysis, LLC, Montvale, NJ

Dr. Frank Gobas, Professor and Chair, Faculty of Applied Sciences, School of Resource and Environmental Management, Simon Fraser University, Burnaby, British Columbia, Canada

Dr. Christian Grue, Associate Professor, Aquatic and Fishery Sciences, Washington Cooperative Fish and Wildlife Research Unit, University of Washington, Seattle, WA

Dr. Charles Hawkins, Professor and Director, Western Center for Monitoring and Assessment of Freshwater Ecosystems, Department of Aquatic, Watershed, and Earth Resources, Utah State University, Logan, UT

Dr. Michael Hooper, Associate Professor of Environmental Toxicology, Institute of Environmental and Human Health, Texas Tech University and TTU Health Sciences Center, Lubbock, TX

Dr. Lynn McCarty, Ecotoxicologist, L.S. McCarty Scientific Research & Consulting, Markham, Ontario, Canada

Dr. Joseph S. Meyer, Professor, Department of Zoology and Physiology, University of Wyoming, Laramie, WY

Dr. Judith L. Meyer, Distinguished Research Professor, Institute of Ecology, University of Georgia, Athens, GA

Dr. Michael C. Newman, Professor of Marine Science, School of Marine Sciences, Virginia Institute of Marine Science, College of William & Mary, Gloucester Point, VA

Mr. Robin Reash, Principal Environmental Scientist, Water and Ecological Resource Services, American Electric Power, Columbus, OH

Dr. Daniel Schlenk, Professor, Department of Environmental Sciences, University of California, Riverside, Riverside, CA

Dr. William Stubblefield, Toxicologist, Parametrix, Albany, OR

Dr. Judith S. Weis, Professor, Department of Biological Sciences, Rutgers University, Newark, NJ

SCIENCE ADVISORY BOARD STAFF

Dr. Thomas Armitage, Designated Federal Officer, U.S. Environmental Protection Agency, Washington, D.C. 20460

Appendix B – Meeting Agenda

SCIENCE ADVISORY BOARD

**Aquatic Life Criteria Guidelines Consultative Panel
SAB Conference Center
1025 F Street, N.W., Suite 3705, Washington, D.C. 20004**

September 21, 2005, Public

AGENDA

- 8:30 - 8:40 a.m. **Meeting Convened by the Designated Federal Officer**
Dr. Thomas Armitage
- 8:40 - 8:45 a.m. **Welcoming Remarks**
Dr. Vanessa Vu, Director, EPA Science Advisory Board
Staff Office
- 8:45 - 8:55 a.m. **Purpose of the Meeting and Review of Agenda**
Dr. Kenneth Dickson, Chair
- 8:55 – 9:20 a.m. **Planned Activities and Overview of Proposed Revision
of EPA’s Aquatic Life Criteria Guidelines**
Dr. Edward Ohanian, Director, Health and Ecological
Criteria Division, EPA Office of Water

Dr. Tala Henry, EPA Office of Water
- 9:20 – 10:00 a.m. **Panel Discussion of Scope of Framework for Revising
Aquatic Life Criteria Guidance**
Lead Discussants: Dr. Charles Hawkins, Dr. Judith Meyer,
Dr. Michael Newman, and Dr. William Stubblefield

Charge Question 1.1: Please comment on the use of the
Guidelines for Ecological Risk Assessment as an essential
and relevant organizing framework for development of
science-based criteria for the protection of aquatic life and
aquatic-dependent wildlife. Does the SAB have any
specific recommendations on how to improve or clarify the
generic conceptual framework diagram?

Charge Question 1.2: Please comment on whether the
proposed criteria types and scientific focus for each criteria

type are logical and scientifically valid for developing a holistic and integrated criteria framework.

10:00 – 10:15

BREAK

10:15 – 10:45 a.m.

Proposed Revisions for Deriving Water-Based Criteria

Mr. Charles Delos, EPA Office of Water

Dr. Russell Erickson, EPA Office of Research and Development

10:45 – 12:15 p.m.

Panel Discussion of Proposed Revisions for Deriving Water-Based Criteria

Lead Discussants: Dr. Joseph Meyer, Dr. Michael Newman, Mr. Robin Reash, and Dr. William Stubblefield

Charge Question 2.1: Please comment on whether the kinetic toxicity models being considered by the EPA are scientifically appropriate for use in deriving water-based criteria.

Charge Question 2.2: Please comment on whether the population models being considered by EPA are scientifically appropriate for use in deriving water-based criteria.

Charge Question 2.3: Please comment on whether the proposal for aggregating effects across species being considered by EPA is scientifically appropriate for use in deriving water-based criteria.

Charge Question 2.4: Please comment on whether the framework being considered by EPA for deriving water-based criteria is scientifically appropriate for use in deriving the criteria.

12:15 – 1:00 p.m.

LUNCH

1:00 – 1:30 p.m.

Public Comments

1:30 – 1:45 p.m.

Proposed Revisions for Deriving Tissue Based Aquatic Life Criteria

Mr. Keith Sappington, EPA Office of Research and Development

1:45 – 2:00 p.m.

Proposed Revisions for Deriving Tissue Based Wildlife Criteria

Dr. Richard Bennett, EPA Office of Research and Development

2:00 – 3:30 p.m.

Panel Discussion of Proposed Revisions for Tissue-Based Criteria

Lead Discussants: Dr. John Connolly, Dr. Frank Gobas, Dr. Christian Grue, Dr. Michael Hooper, and Dr. Lynn McCarty

Charge Question 3.1: Please comment on the rationale and conceptual approach used for the development of tissue-based criteria for this group of chemicals. Is the SAB aware of other approaches for deriving criteria for these bioaccumulative chemicals that EPA should consider?

Charge Question 3.2: Considering the strengths and limitations of the more flexible approach used to derive tissue-based criteria, please comment on the rationale and preference for allowing flexibility in the procedures used.

Charge Question 3.3: Please comment on the rationale used by EPA for determining if/when to use population modeling in the development of Tissue-Based criteria.

3:30 – 3:45

BREAK

3:45 – 4:00 p.m.

Proposed Revisions for Deriving Taxon-Specific Criteria

Mr. Brian Thompson, EPA Region 5
Mr. Thomas Augspurger, U.S. Fish and Wildlife Service

4:00 – 5:00 p.m.

Panel Discussion of Proposed Revisions for Taxon-Specific Criteria

Lead Discussants: Dr. Charles Hawkins, Dr. Daniel Schlenk, and Dr. Judith Weis

Charge Question 4.1: Please comment on the considerations for problem formulation outlined in the proposed framework for deriving Taxon-specific Criteria, specifically whether it will lead to scientifically-defensible numeric criteria.

Charge Question 4.2: Of the approaches outlined for addressing surrogacy and gap analyses with regard to

special status species, are there improvements to these tools that would provide more scientifically-defensible numeric criteria where specific data are not available? Are these tools adequate for developing scientifically-defensible numeric criteria? What other tools are available to provide more scientifically-defensible criteria when there is an absence of toxicological data for a specific pollutant and taxon?

5:00 – 5:15 p.m.

Summary of the Discussion
Dr. Kenneth Dickson, Chair

5:15 p.m.

Adjourn

Appendix C – Initial Responses to Charge Questions

Initial Response to the Charge Questions Pertaining to Tissue-Based Criteria

John P. Connolly
Quantitative Environmental Analysis, LLC

For chemicals with a high propensity to bioaccumulate in aquatic food webs and for which diet is a primary route of exposure, the EPA proposes to develop tissue-based criteria expressed as the chemical concentrations in specific animal tissues or dietary concentrations, with a process for translating to corresponding water and sediment concentrations. Tissue-based criteria allow for integration of multiple exposure pathways (water, diet) and facilitate direct comparison with environmental tissue concentrations to determine if there is a risk of adverse effects.

Charge Question 3.1. Please comment on the rationale and conceptual approach used for the development of tissue-based criteria for this group of chemicals. Is the SAB aware of other approaches for deriving criteria for these bioaccumulative chemicals that EPA should consider?

Tissue-based criteria have great appeal for the reasons presented in the Science Advisory Board Consultation Document (SABCD). As I read this listing of reasons, I noted that the arguments for the use of tissue-based criteria that are presented in Section 2.2 of the SABCD apply to a much broader range of chemicals than those for which multiple routes of exposure are important. The toxicokinetic differences among species and individuals noted in Section 2.2 (i.e., differing rates of uptake, distribution, metabolism, and elimination) are confounding factors even for chemicals whose uptake is principally via a single route. Moreover, regardless of whether multiple routes of exposure are important, tissue-based measures of toxicity are not subject to the factors that affect chemical bioavailability in a laboratory toxicity test and confound the interpretation and grouping of media-based toxicity data.

The SABCD notes that a tissue-based approach is being developed for cationic metals (i.e., the Biotic Ligand Model) and the publications cited in the last paragraph of Section 2.2 to support the use of a tissue-based approach describe the application of a tissue-based approach to a broad range of chemicals that is not restricted to those for which multiple routes of exposure are important. Thus, while the rationale for using tissue-based criteria is clearly articulated, I perceive a need to better articulate the reason for restricting the use of such criteria to chemicals “with a high propensity to bioaccumulate.” I think that such articulation is of particular importance because conversion of a tissue-based criterion to a media-based criterion, a step in the criteria development process that I

view as a major challenge to the effective use of tissue-based criteria, is simpler for many chemicals with a single route of exposure than for the class of chemicals for which tissue-based criteria are proposed. Thus, I believe it is incumbent on the Tissue-based Criteria Subcommittee to make the case for limiting the scope of tissue-based criteria to bioaccumulative chemicals. This may be as simple as discussing the lack of tissue-based toxicity data, but there may be other reasons known to the Subcommittee.

The conceptual approach for developing tissue-based criteria seems to be more complete for the consideration of toxicity data than for the translation from a tissue-based criterion to a media-based criterion. The approach acknowledges the key issues that must be addressed to develop a defensible tissue-based criterion, but it does not provide the same level of provisos for the translation step. In my view, the road blocks to translation are substantial and include, among other things, the following:

- Spatial variability in exposure – the current practice for bioaccumulative chemicals assumes that the animals live in the discharge pipe (i.e., there is no mixing zone or allowance for animal movements)
- Appropriate time averaging of the media concentration – presumably this would depend on the nature of the tissue-based concentration (e.g., species and life stage)
- Relationship between sediment and water column media concentrations – how are the two media correlated?

I suggest that the conceptual approach for translation be expanded to address the key issues with the degree of comprehensiveness provided for the derivation of the tissue-based criterion.

The idea of using a steady-state approach for tissue-based criteria is reasonable, but the SABCD discussion of what this means with regard to interpretation of a criterion value appears to be limited to an acknowledgement of the issue in Section 3.1.2. I think that the Sub-committee should give consideration to the development of specific guidance regarding how one interprets a long-term average criterion concentration in the real world where concentrations vary greatly in space and time. How does one establish an appropriate averaging period and spatial scale for averaging? Should growing season be a factor?

A Minimum Data Requirement (MDR) for deriving deterministically-based criteria of approximately 4-5 species is suggested. This seems like a reasonable minimum to achieve criteria that have some minimum level of realism. However, I was not able to judge from the data summary presented as Appendix A whether this requirement would effectively preclude the development of tissue-based criteria pending the completion of a substantial number of new toxicity studies. It would be helpful if the appendix presented a table showing, by assemblage (i.e., vertebrates, invertebrates, plants), the number of chemicals for which toxicity data of appropriate exposure duration, a common class of effect and common tissue type exist for at least 4-5 species. It would also be helpful if an indication was made of whether a tissue concentration-response relationship had been

demonstrated for that tissue type. Such a data summary might help determine whether it is practical to proceed with the development of tissue-based criteria and, if not, the level of effort and time necessary to generate the required data. Are tissue-based criteria practical now, in five years, or even in ten years?

The proposed process for Tissue-based Criteria is intended to be flexible to maximize the use of available data and to accommodate certain limitations in the quality and quantity of data. This approach will also provide opportunities for states and tribes to develop alternative options that may be more suitable to site-specific conditions. National-level criteria may use deterministic approaches to characterize toxicity data when data are limited or probabilistic approaches (e.g., species sensitivity distributions) when data are sufficient. The process will also describe how a criterion may be refined on a site-specific basis when additional data are available.

Charge Question 3.2. Considering the strengths and limitations of the more flexible approach used to derive tissue-based criteria, please comment on the rationale and preference for allowing flexibility in the procedures used?

I agree with the idea of using an assessment of the quantity and quality of data to determine whether a tissue-based criterion can be developed and if it should be developed using a deterministic-based approach or a probabilistic approach. The Sub-committee clearly recognizes the limitations of existing data and its proposed flexible approach does provide a potential means to develop tissue-based criteria despite those limitations. However, the allowance for flexibility in the treatment of toxicity data and in the conversion of a tissue-based criterion to a media-based criterion has the potential to substantially corrupt the final value and defeat the advantages of tissue-based criteria laid out in Section 2.2. I suggest that the Sub-committee consider the extent to which the uncertainties introduced in an effort to achieve the MDR compromise the reductions in uncertainty achieved by using tissue-based criteria. It seems to me particularly important that extrapolation, as described in Section 3.2.2, be done cautiously and with an explicit accounting of uncertainty. In my view, the resulting uncertainty estimates should not be used to drive a profoundly conservative criterion; they should be used to judge whether the uncertainty of a tissue-based criterion is such that a tissue-based criterion is inferior to a water-based criterion. The potential economic impacts of water quality criteria are enormous and I believe that it is essential that the Sub-committee ensure that the relative accuracy and uncertainty of water-based and tissue-based criteria be evaluated and that the more accurate and least uncertain value be the one chosen as a criterion.

The translation of tissue criteria to concentrations in water and the food web is flexible of necessity given the highly site-specific nature of bioaccumulation and the absence of site-specific data for some chemicals and many sites. I agree in general with the stated preference to use site-specific data where available and a model where such data are not available. However, the devil is *really* in the details on this point. I urge the Sub-committee to develop comprehensive procedures for the compilation and use of data and for the development of a model. It has been my experience that appropriate pairing of biota and media data is a non-trivial exercise requiring an in-depth understanding of the

spatial and temporal distributions of concentration and the life history of the biota. At my favorite site (the Upper Hudson River), which may have the greatest density of data of any site in the world, the relationships between biota, water and sediment PCB concentrations vary greatly depending on location within the river and the means by which data are aggregated in time and space. A realistic understanding of bioaccumulation at this site has required the combined use of the comprehensive existing data set and bioaccumulation modeling. As with the use of toxicity data, the uncertainty of translating tissue concentration to water and food web concentrations should be an important consideration in the development of tissue-based criteria. Unlike the dynamic exposure scenarios being addressed in development of water-based criteria, EPA is considering a steady-state approach for developing national criteria for bioaccumulative chemicals (i.e., modeling bioaccumulation and toxicity as a function of constant concentrations). Rationale for this approach is the much slower accumulation kinetics generally associated with these chemicals in higher trophic level fish and aquatic-dependent wildlife and concerns over their long-term bioaccumulation. In the context of population modeling, there appears to be much less residue-response information available for integrating responses of various demographic parameters over multiple life stages, such as fecundity and adult, juvenile, and larval survival. Consequently, it is not clear whether it would be feasible or useful to integrate population modeling into national-level tissue criteria for bioaccumulative chemicals. Current thinking is that where sufficient data exist to characterize exposure, bioaccumulation and toxicity on a dynamic basis, population modeling may evolve into an important tool in the development of site-specific criteria.

Charge Question 3.3. Please comment on the rationale used by EPA for determining if/when to use population modeling in the development of Tissue-Based Criteria?

I struggle to reconcile the arguments against population modeling with the arguments for the development of tissue-based criteria. If the effect of chemicals on populations is driven by fluctuating concentrations and particular life stages, what is the meaning of a criterion that protects a default life stage defined by data availability against the long-term average concentration to which it is exposed? Having said this, I do not disagree with the decision to defer any consideration of population modeling because of the lack of sufficient data to conduct meaningful simulations. I suggest that the Sub-committee consider whether my difficulty in reconciling the approach to tissue-based criteria with the arguments against population modeling indicates a weakness in the underpinnings of the approach or a misinterpretation on my part.

Initial Responses: Frank Gobas

Please, comment on the rationale and conceptual approach for the development of tissue based criteria for this group of chemicals. Is the SAB aware of other approaches for deriving criteria for these bioaccumulative chemicals that the EPA should consider?

The process of using tissue based criteria for the development of water quality and sediment quality objectives for bioaccumulative compounds is fundamentally a sound approach which implicitly recognizes and correctly treats the concepts of exposure and potency/toxicity. In my view the tissue based approach discussed in this document should not be limited to “bioaccumulative” substances but applied universally, i.e. to less bioaccumulative substances as well, because it is the only approach that recognizes the differences in the water-internal concentration relationships that exist between laboratory based toxicological tests and field situations. In terms of the determination for the rationale for a tissue-based criteria approach (p.12, p.16), I therefore recommend adding to the rationale the criterion related to differences between water- internal concentration relationships between laboratory based toxicological tests and field situations (e.g., resulting from short exposure duration in test, bioavailability differences between lab and field, experimental problems maintaining water concentrations in tests, differences in tissue composition (e.g. lipid content) between lab and field animals, metabolism /induction).

I recommend caution on the application of Kow as a property to distinguish between so called “bioaccumulative” and “non-bioaccumulative substances”. A Kow of 5 is suggested in the document in various places. There are two major reasons: First, disequilibria between suspended matter and water and also bottom sediment and water appear to be greatest for low Kow chemicals. This means that dietary consumption of particulate matter can be a significant exposure route even for low Kow chemicals as the actual particle –water distribution coefficients in the field are much greater than anticipated based on Kow based equilibrium partitioning. Secondly, while Kow is an appropriate parameter to indicate food-web magnification in water-breathing organisms, it is not indicative of the biomagnification potential of chemicals in air-breathing organisms. It should be recognized that poorly metabolizable chemicals with a log Kow > 2 and a log Koa > 5 have been observed to be bioaccumulative in air-breathers. I have serious reservations about the application of the species sensitivity distribution (SSD). The current proposal suggests the compilation of effects based tissue residue concentrations (for bioaccumulative substances) or dietary concentrations (for wildlife) and the selection of an appropriate percentile from the effects distribution. The application of such an approach has some theoretical and pragmatic limitations. First, any selected percentile that is statistically defensible (e.g. 95% mentioned in several places) is ecologically unjustifiable, even in the rare cases where a large number of data is available. In the application of this approach we have to be aware of a key assumption that is being introduced, namely that the percentile of test results showing an effect is de facto assumed to be equal to the percentile of species / individual organisms (?) affected.

A 5% effect can be an ecological catastrophe. Applying additional uncertainty factors may be useful in this regard. The merit of the uncertainty factors is to implicitly recognize uncertainty in the application of test result distributions to real-world species distributions.

Secondly, the national tissue concentration that would be adopted based on the SSD is for a somewhat ‘generic’ but non-existent organism. It is more scientifically defensible to make use of the available species specific toxicity data base, e.g. use rainbow trout effects data to derive a rainbow trout specific tissue residue criterion which is then applied to rainbow trout to calculate ecosystem specific water and sediment concentrations to protect rainbow trout. For species for which similar effects data do not exist, a probabilistic and deterministic SSD can be used.

While defining bioaccumulation behavior in representative species is useful in cases where a site-specific approach cannot be carried out, I recommend pointing out on p. 14 one of the main strengths of the tissue criterion approach, namely to derive water and sediment quality guidelines that are ecosystem specific. The possibility of a site specific approach will increase the confidence of stake holders in the approach and the need for any remediation type activities that may follow. My suggestion is to add in the boxes on the bottom of Fig. 1 “site-specific” species.

I strongly agree with the approach where aquatic dependent wildlife (e.g. mammals, birds, reptiles) are included in the derivation of national tissue concentrations. However, I do not recommend that this is done based on dietary concentrations as is proposed (p. 13) and provided as an option on p.34. The reason is that diet- internal tissue concentration relationships in test animals can vary substantially from those in the field, for many of the same reasons that led to the proposal for using tissue-residue based criteria, e.g. differences in exposure duration between lab and field, chemical bioavailability, differences in animal tissue composition (e.g. lipid content). I recommend that the tissue concentration approach is also exclusively used for wildlife. This can be done by using toxicokinetics models. It would eliminate the need to develop UFL and UFs (p. 40), which relate to largely known differences in experimental design (e.g. exposure duration) rather than true uncertainty (e.g. in inter species sensitivity). Toxicokinetics factors may also be responsible for a significant part of the two orders of magnitude differences in toxicity among wildlife species.

In terms of one of the key challenges to a tissue concentration approach, namely access to tissue-concentration based toxicological measurements (p.15), I recommend that the Agency further explores the application of toxicokinetics models and internal pharmacokinetic models in compiling effects based tissue concentrations from water and diet based toxicity data. Application of these models will allow for a use of more toxicological data. It will also address problems noted on p.20, “exposure duration can contribute to variance in tissue-based toxicological effects levels.” And “chemical concentrations in some tissues having little or no correlation with toxicological effects”. In my view, it is therefore crucial to apply toxicokinetic and pharmacokinetic models to derive appropriate tissue-concentration-response relationships in tox studies.

The bioaccumulation models that are referred to in the proposal to derive water and sediment based criteria from the tissue concentrations can be also be used to back calculate tissue concentrations from water or diet based toxicity data.

In my view, the Agency is better off (in terms of scientific defensibility, transparency and amount of work) translating non-tissue concentration based toxicological data (i.e. toxicological data in terms of water concentrations, sediment concentration, dietary concentrations) and exposure duration linked toxicological data (chronic and non-chronic exposures) in terms of tissue concentrations, rather than developing conceptually different approaches for wildlife (dietary concentrations) vs. fish/invertebrates (tissue concentrations) and bioaccumulative (based on tissue concentrations) vs. non bioaccumulative substances (based on water concentrations). In the case of wildlife, this would eliminate the need to develop UFL and UFs (p. 40), which relate to largely known differences in experimental design (e.g. exposure duration) rather than true uncertainty (e.g., in inter species sensitivity).

p.21 For many hydrophobic organic chemicals, the lipid normalized concentration is a better alternative than whole body concentrations in the derivation of a tissue concentration response relationship. It may therefore be advantageous to express tissue residue concentrations on a lipid normalized basis for many hydrophobic organic substances.

In the section on characterizing effects on organisms, there is a lack of discussion on nonthreshold effects, such as tumor incidence, immunotoxic responses. Although I think that the tissue residue approach can be used for these effects as well, there is a need to consider methods for high-dose to low-dose extrapolations.

In terms of deriving a tissue residue criterion from a variety of toxicological endpoints, has Subcommittee developed any criteria to what toxicological effects should be included or excluded or weighted more than others? For example, it is possible that for substance A only acute mortality data are available while for another substance B less acute mortality data are available in addition to the acute mortality data. All else being equal, will the current approach derive a more stringent criterion for chemical B compared to A?

p.24. I do not fully agree that extrapolation from EC50 to EC10 or LOAEL to NOAEL involves an extrapolation across the magnitude of effect. The EC10 refers to the same effect as the EC50. The difference between EC10 and EC50 is a measure of the uncertainty in the characterization of only one effect, i.e. 50% mortality. This uncertainty is important to know for establishing a SSD, but it should not be used to characterize a lower level of effect.

p.43. For wildlife feeding on multiple trophic levels, it is not clear to me how tissue concentrations in representative organisms of different trophic level can be derived given that multiple combinations of concentrations in dietary items of different trophic level can give the same concentration in the wildlife consumer species.

Considering the strengths and limitations of the more flexible approach used to derive tissue -based criteria, please comment on the rationale and preference allowing flexibility in the procedures used.

As for the approach's flexibility in terms of using water concentration, tissue concentrations and dietary concentrations as the basis for criterion development, I think that this is not a strength but a weakness for several reasons:

1. It is much more complicated as there are now 3 approaches and it is unclear which substance should be derived with what methodology.
2. The main strength of the tissue concentration approach is that it recognizes that a toxicological response is controlled by two processes, i.e. exposure (the relationship between external concentration and internal concentration) and toxicity (i.e. the internal concentration associated with an effect). The toxicity part of the equation is in many cases a common quantity in the tox text and the real world. The exposure part (i.e. the relationship between external and internal concentration) varies widely between the lab and field but in a way that to a large degree is understood. By not distinguishing between exposure and toxicity, such as is the case in the water based criteria and diet based criteria, largely known toxicokinetics processes and factors affecting a toxic response are treated as uncertainty (e.g. UFL and UFs (p. 40), leading to an apparent high degree of uncertainty (as uncertainty factors are multiplied to achieve high values) while much of this uncertainty does not exist.
3. The flexibility also divides the toxicological data base in subsets that are evaluated independently (to derive an SSD) whereas there is a lot of value in evaluating the toxicological data base as a whole.

I strongly recommend that the Agency develop a tissue concentration data base that includes toxicological data expressed in terms water and dietary concentrations in addition to those expressed as internal tissue concentrations. As for flexibility in terms of a site specific and a representative species approach, I think that this is a good idea.

Please comment on the rationale used by the Tissue based criteria subcommittee for determining if/when to use population modeling in the development of tissue based criteria.

Although theoretically I do not see a good reason for why populations models could not be used for deriving national criteria, there are some substantial practical problems associated with doing this. The document discusses some. Another important problem with the use of population models is that any effect that the chemical exerts has to be considered in the context of all other stresses that act upon the population (e.g. predation, habitat destruction, hunting, fishing, temperature, water levels, disease, etc, etc, etc.). Currently, the state of the science is not sufficient to include this additional realm of complexity. In many cases there may not even to be a need for this, e.g. in cases where

human consumption controls the tissue residue criterion. Also, the current approach which focuses on protecting the individual will ensure that the population is protected as well.

Science Advisory Board
Aquatic Life Criteria Guidelines Panel Meeting
21 September 2005

Comments from Christian Grue

“Tissue-based Criteria for “Bioaccumulative” Chemicals

General comments

The Agency should be commended for its efforts to address recommendations outlined in the 2003 report summarizing discussions between 1985 and 1995. Much time has passed since 1995. The Agency notes that the development of revised criteria has been hindered by ESA consultations – but no details are provided. I assume the issues are captured in the proposed guidance revisions.

It would be helpful for the Agency to describe how existing water quality criteria and those generated from future guidance will be applied. Having recently worked with a document in which existing acute and chronic water quality criteria were applied to concentrations of pesticides detected in grab samples from surface waters at different points in time, I question whether or not they are being used/interpreted correctly. To what extent the sampling designs being employed in the field appropriate for the application of these criteria? Guidance from the Agency in applying these criteria will be even more important if additional types of criteria are developed. It is also not clear to what extent State’s have discretion in using or modifying national criteria.

Within the context of the proposed guidance revisions, it is surprising that a greater emphasis has not been placed on “non-traditional” endpoints, specifically “sublethal” effects such chemical-induced changes in behavior which can decrease survival, interfere with reproduction, and lead to changes in community structure at concentrations much lower than those associated with direct mortality of adults or embryos. At one extreme are overt changes in behavior such as immobilization and narcosis, and at the other changes in physiology altering reproductive behavior or disease resistance. Additionally, an increasing number of studies are showing that exposure to additional stressors can cause “sublethal” exposures to become lethal.

Similarly, one can argue that without a good understanding of the factors governing exposure in the field, it will be difficult to predict what an individual’s actual exposure will be, particularly if choices (gradients in water concentrations or prey) exist that can be dictated by abiotic and biotic factors (i.e., behavioral modification of exposure). One could also argue that tissue criteria may be more appropriate than water criteria because of the better integration of these variables.

The toxicity associated with chemical mixtures also continues to be under emphasized or ignored. Recent studies have reinforced earlier studies indicating that active ingredients may not be driving the toxicity of formulated products and tank mixes and that the

“other” ingredients warrant increased scrutiny at the federal or state level. Furthermore, as an increasing number of studies report chemical concentrations in surface waters, efforts to interpret these data will force an examination of chemical mixtures. I would argue that addressing this issue by only examining chemicals with the same mode of action is not adequate.

Many of these comments reinforce the potential need for more local criteria, or at least the guidance from which more local criteria can be developed. The extent to which states (or other federal agencies) will develop their own criteria is unclear, but in part will depend on the quality of the data and analyses from which the national criteria are developed.

Availability of appropriate data continues to be problematic. How is the Agency planning to address these data gaps? We need a federal initiative (funding) to provide the necessary data. Without it, our progress will be severely limited. In addition, the number of assumptions supporting extrapolation and modeling and the magnitude of safety factors will also increase. The later need to be clearly defined and supported.

It would be helpful to know how Canada and the European Union are dealing with these issues as we hopefully move toward harmonization of regulatory requirements for data and testing.

Charge Questions

3.1. Please comment on the rationale and conceptual approach used for the development of tissue based criteria for bioaccumulative chemicals? Is the SAB aware of other approaches for deriving criteria for these chemicals that EPA should consider?

I am not convinced the working definition of “bioaccumulative chemicals” needs to be as stringent as proposed. Bioaccumulation can be defined as the uptake of chemicals by water and food. With respect to species of aquatic-dependent wildlife, biomagnification may be a better term, as the primary route of exposure is through the diet and concentrations of the chemicals described in these species will undoubtedly be greater than species at lower trophic levels.

I agree with the rationale for developing tissue based criteria, the application of the risk assessment framework, and support the harmonization of proposed tissue based criteria guidance with bioaccumulation assessment guidance for human health. If data and analyses are sufficient to protect human health, they should be adequate to protect aquatic resources. I don’t think the use of dietary concentrations for aquatic-dependent wildlife is problematic and is justified based on the availability of data and the greater ease of collecting relevant field data (concentrations in food items) in the future. Sensitivities

associated with sacrificing aquatic-dependent wildlife to obtain tissue residues may be problematic

As with any of these approaches, a primary concern in terms of testing, and ultimately implementation, is data availability. The data review described in Appendix A raises concerns. Specifically, the usable data available are probably “significantly less” because data have not been screened, nearly all of the chemicals are represented by 5 or fewer species; few chemicals include data on 5 or more species; mortality is the predominant endpoint; whole bodies are the principal tissue represented; water is the primary route of exposure; life stages are not equally represented, primarily adults and juveniles; and most of the data are for salmonids. Potentially more problematic is the lack of cause-effect data, i.e. modes of action and target tissues, although the use of whole body concentrations (adjusted on a lipid weight basis?) as a surrogate for specific tissues needs to be examined. EPA needs to describe the methods it will use to predict concentrations in specific tissues from whole body data. This is not a problem for aquatic-dependent wildlife because, in most cases, whole body residues will be most appropriate.

As with the two other types of criteria proposed (water and taxon), I suggest EPA use the available data to generate specific examples of the conceptual approach proposed. It appears sufficient data are available to provide an example of deterministic and probabilistic calculations. As the Agency notes, population modeling may not be possible because of the few data for specific life stages.

3.2 Considering the strengths and limitations of the more flexible approach to derive tissue based criteria, please comment on the rationale and preference for allowing flexibility in the procedures used?

The rationale for the flexibility in the proposed analytical strategies, deterministic vs probabilistic, is driven by the current availability of appropriate data. While this flexibility may be needed at the present time, EPA is encouraged to develop the data appropriate for moving toward probabilistic assessments. I don't see any other options.

3.3. Please comment on the rationale used by EPA for determining if and when to use population modeling in the development of tissue based criteria?

The rationale for using, at least initially, a steady state model (constant chronic exposures) for developing tissue criteria for bioaccumulative chemicals vs the dynamic model used to represent daily fluctuations in water concentrations to develop water based criteria for other chemicals makes sense. However, as the Agency notes, species for which the kinetics for the subject chemicals vary more rapidly, will need to be treated differently.

With respect to population modeling, again an apparent lack of life stage specific data appears to be problematic.

Initial Comments from Charles Hawkins

- 1.1 Use of the Guidelines for Ecological Risk Assessment as an essential and relevant organizing framework for development of science-based criteria.
 1. The broadened recognition of a need for different types of criteria is a critical improvement over the old approach. EPA needs to provide more details regarding how multiple criteria will be applied in practice when more than one criterion is applicable.
 2. The materials presented were reasonably free of excessive jargon, which is often mind-numbing in EPA documents.
 3. I found the conceptual framework extremely useful, especially as a means of describing risk hypotheses and the attendant assumptions associated with hypotheses.
 4. The conceptual framework needs a bit of tweaking to be explicitly clear. For example, I do not understand what “translation of chemical concentrations among compartments” means in the context of the bi-directional arrows. I also don’t know what differences are implied by the dotted and solid lines.
 5. The main limitation of the framework is not its general approach but on its apparent restricted use in addressing largely (only?) Chemical criteria (see 1.2 below).
- 1.2 Are the criteria types and scientific focus for each criteria type logical and scientifically valid (for developing a holistic and integrated criteria framework).
 1. I thought the 3 types of criteria described were logical and valid and probably generally applicable to chemical pollutants.
 2. However, I question whether the framework is either truly holistic or integrated in that there is essentially no discussion of how the framework could (or will) be applied to non-chemical stressors that might be best addressed through the development of “habitat-based” criteria. For example, main causes of degraded aquatic life in the Nation’s streams and rivers are the direct and indirect effects of habitat destruction, altered flow regimes, excesses sediment above natural levels, thermal alteration, and excessive nutrient loads. I understand the legacy of chemical criteria within EPA and that this new framework builds on a simpler and less realistic one, but it still touches the tip of an iceberg. It seems to me we need to think about a common framework that works for the whole iceberg and not just chemical and related pollutants. Failing to do so now can only cause problems with later required revisions designed to address this issue.
- 4.1 Is the approach to problem formulation in developing taxon-specific criteria defensible?

1. Overall, I had few comments regarding the general approach to problem formulation. It appears to be a logical and straight-forward approach.
 2. This document did acknowledge that factors other than chemistry need to be considered when deriving criteria. To the extent that species of special concern are often faced with loss of critical habitat, the overall framework (1.1-1.2) is not explicit in how such criteria would be developed.
- 4.2 Are there improvements to these tools that would provide more scientifically defensible numeric criteria where specific data are not available?
1. I suspect the ICE estimates might be improved by following a somewhat different statistical logic than presented. This is largely a detail issue, but I was unconvinced that the criterion recommended for deciding which estimate to use when several estimates were available was the most robust approach. If all estimates are independent, the mean of these estimates should be closest to the true value, no?
 2. I also think that the adequacy of many of the possible specific tools will require scrutiny from several ecologically-oriented statisticians. In some instances, understanding the adequacy of the tools discussed in the report will require familiarity and experience with the techniques described as well as a bit more formal knowledge than I possess. What seems sensible on the surface can sometimes be fraught with problems, i.e., devil is in the details.

Initial Responses to the Aquatic Life Criteria Guidelines Panel Charge Questions

Michael J. Hooper
Texas Tech University

Charge Question 1.1

Please comment on the use of the Guidelines for Ecological Risk Assessment as an essential and relevant organizing framework for development of science-based criteria for the protection of aquatic life and aquatic-dependent wildlife. Does the SAB have any specific recommendations on how to improve or clarify the generic conceptual framework diagram?

The outlined approach of using a modified version of the GERA, i.e., progressing from problem formulation through effects characterization and using risk characterization process to “back-calculate” exposure limits appears sound. The application of a detailed problem formulation process for both the establishment of criteria development guidelines as well as the subsequent specific criteria themselves will provide clarity and transparency to the process.

Charge Question 1.2.

Please comment on whether the proposed criteria types and the scientific focus for each criteria type are logical and scientifically valid for developing a holistic and integrated criteria framework.

The general breakdown of criteria types into water, tissue and taxon specific approaches is appropriate for the development of AWQCs.

“Tissue-Based Criteria” Questions

Charge Question 3.1.

Please comment on the rationale and conceptual approach used for the development of tissue-based criteria for this group of chemicals. Is the SAB aware of other approaches for deriving criteria for these bioaccumulative chemicals that EPA should consider?

1. Working with wildlife monitoring programs over the years, one of the most frequent laments has been the lack of tissue residue data for animals suspected to have died or been impacted by chemical contaminants. Dose-response studies that identify toxic effects with dose generally fail to measure the levels of chemicals that build up in tissues as intoxication proceeds from normal health, through subtle sub-lethal effects to incapacitation and death. The use of contaminant tissue levels to assess hazard to wildlife is a desirable approach that provides a measure more relevant to the health of the protected species. Unfortunately, as the text puts it adeptly (pg 23),

Perhaps the greatest obstacle facing the successful derivation of tissue-based aquatic life criteria for bioaccumulative chemicals is the relative lack of appropriate, standardized, tissue-based toxicological data.

Given a move to these types of methods, a regulatory mechanism for generating the necessary data is needed. Much of this data might come with studies carried out concurrently with chronic or reproductive toxicity tests.

2. The use of dietary item chemical levels is generally less successful for wildlife, because of their feeding in heterogeneously contaminated habitats compared to aquatic species that are found together with the majority of their food chain in the same body of water. Wildlife tissue data provide a more integrated estimate of exposure, particularly for those species that might travel between a variety of water bodies within their home range.
3. Though an understanding of mechanism of action of chemical contaminants is useful and to be encouraged, there are some contaminants that are best assessed by measuring contaminants in a tissue where residues best predict level of toxic effect, without a particularly strong rationale. Brain dieldrin levels are particularly useful for predicting dieldrin toxicity across a wide variety of species due to the effect of dieldrin on brain GABA receptors. Alternatively, Pb accumulation in the liver provides a similar degree of toxicity predictability, though the effects of this metal are diffuse across many tissues and systems.
4. Along similar lines, consideration of mechanisms of action and associated tissue contaminant levels for sub-lethal endpoints that can have adverse effects on wildlife, including immunological suppression and endocrine modulation, for example, would provide more sensitive assessment endpoints
5. Though probabilistic methods have lately been shown to have greater utility in toxicity-associated risk assessments, the necessity of having a distribution of chronic or sub-chronic study data for particular species of wildlife is likely to decrease the likelihood of their use in criteria establishment. Aside from a multitude of avian species studied for impacts of DDT, there are likely few other chemicals for which sufficient chronic/sub-chronic data are available to develop a distribution of species.
6. It is worth noting in the section on extrapolation that QSAR models might help in developing criteria for a family of similar structural compounds if the approach can be validated appropriately.
7. Though the additivity of multiple chemicals with similar mechanisms of action was mentioned in the introductory comments, note should be made of chemicals that are metabolized into a number of active metabolites with different dose-response relationships in the exposed animal. Criteria for metabolized compounds should be developed using values weighted to account for this differential toxicity.

Charge Question 3.2.

Considering the strengths and limitations of the more flexible approach used to derive tissue-based criteria, please comment on the rationale and preference for allowing flexibility in the procedures used?

Flexibility is a requirement for tissue-based assessments, as there is substantial variability in the availability of useful tissue contaminant data for wildlife. This approach is appropriate for criteria development.

Charge Question 3.3.

Please comment on the rationale used by EPA for determining if/when to use population modeling in the development of Tissue-Based Criteria?

For wildlife species, I believe that it is going to be enough of an exercise to develop individual-based criteria, I have my doubts that it is desirable at this point to develop population models.

There is, however, a likely role for models of community interaction. They would provide a mechanism for dealing with the issue of indirect effects, an important and generally overlooked effect of environmental contamination. Concern over breeding duck success in a contaminated pond might suggest the need for an avian criteria value. If that contaminant eliminated a particular snail upon which ducklings depend heavily, snail tissue contaminant levels could provide a more sensitive criteria that would be protective of ducks.

**Responses to Charge Questions Posed to the
USEPA SAB Aquatic Life Criteria Consultative Panel**

Prepared by
Dr. Joseph S. Meyer, Professor
Department of Zoology and Physiology
University of Wyoming

16 September 2005

Charge Question 1.1: Please comment on the use of the *Guidelines for Ecological Risk Assessment* as an essential and relevant organizing framework for development of science-based criteria for the protection of aquatic life and aquatic-dependent wildlife. Does the SAB have any specific recommendations on how to improve or clarify the generic conceptual framework diagram?

The use of the *Guidelines for Ecological Risk Assessment* as a general organizing framework appears to be appropriate. But because the *Guidelines* are so general, I will be more interested in the details of how the *Guidelines* will be interpreted during the criteria-revision process.

I have no recommendations on how to improve or clarify the generic conceptual framework diagram.

Charge Question 1.2: Please comment on whether the proposed criteria types and the scientific focus for each criteria type are logical and scientifically valid for developing a holistic and integrated criteria framework.

I am pleased to see that the potential importance of dietborne-contaminant effects might be incorporated into criteria. Acknowledging the importance of taxon-specific concerns is also encouraging, although I'm not sure why the method for specific taxa is isolated from the water-based criteria and tissue-based criteria methods. For example, I don't understand why consideration of specific taxa of concern couldn't be included in the water-based criteria method, with the possibility that the sensitivity of a specific taxon of concern (determined from modeled population trends) could trump the otherwise acceptable toxicant concentration determined from analysis of the population trends within the appropriate species assemblage.

I am also concerned that a sediment-based criteria method is not being concurrently proposed to dovetail with the water-based and tissue-based criteria methods. If the EPA's goal is to develop a holistic and integrated criteria framework, the concurrent exposures to a contaminant in water, sediment, and food should be considered inseparable in aquatic systems.

Additionally, Schnoor et al. (1997) recommended adopting a watershed approach for metals criteria almost a decade ago. I was surprised to see no discussion of the advantages and disadvantages of that type of approach in these review documents.

After stating all that, I am concerned that the lofty goals proposed by EPA are beyond current reach with the available data and some of the current scientific understanding. I applaud EPA for trying to expand the aquatic-life criteria envelope with state-of-the-art science, but it might take decades and considerable funding to develop datasets and additional understanding appropriate to allow some of the proposed methods to be used routinely. Therefore, I am interested in how EPA plans to revise the aquatic criteria methods with available and easily obtainable data and scientific understanding in the interim.

Charge Question 2.1: Please comment on whether the kinetic toxicity models being considered by EPA are scientifically appropriate for use in deriving water-based criteria.

As long as high accuracy is not required, kinetic toxicity models appear to have satisfactorily predicted the survival of aquatic organisms exposed to time-varying concentrations of the relatively few contaminants with which they have been tested. However, to my knowledge, no one has tested the ability of kinetic models to predict effects of time-varying concentrations of contaminants on reproduction and age of first reproduction (the two other primary parameters besides survival that are needed in population models) of aquatic organisms.

Regarding a preference for one kinetic toxicity model over the other, I have not seen any head-to-head comparisons of the deterministic and the stochastic process models that would indicate a clear difference in their abilities to predict the toxicity of time-varying concentrations of contaminants. This is an important comparison that could be accomplished with currently available datasets.

Charge Question 2.2: Please comment on whether the population models being considered by EPA are scientifically appropriate for use in deriving water-based criteria.

The population models being considered by EPA are scientifically appropriate. I am pleased to see that EPA is attempting to incorporate density-dependent population projections into aquatic life criteria. However, it is debatable whether the appropriate baseline condition to assume for most populations of fish, invertebrates and algae is carrying capacity. Even in the absence of contaminants, many populations in aquatic systems probably are well below carrying capacity because of disease, predation, etc. Therefore, compensatory responses in real-world populations might be much less important than in population models that start at carrying capacity (i.e., at the 100% population level, as appears to be planned for the revised criteria approach).

Moreover, I am concerned the data requirements for density-dependent modeling will be a major constraint, because three-dimensional response surfaces will have to be

available for each parameter of interest (e.g., survival in each life stage, growth (or some other index of development) in each pre-adult life stage; age of first reproduction; and fecundity of adults) as a function of a gradient of contaminant concentration and a gradient of population density (up to and exceeding carrying capacity) -- for each species included in the calculation of a criterion. Not enough appropriate data are available to currently include many (if any) species in this type of density-dependent criteria calculation, and I suspect development of robust datasets will take many years and consume considerable funding.

Regarding the choice of a stage-based matrix model for population projections, I have two concerns that appear to be challenging but not necessarily insurmountable.

1. Stage-based models usually are incremented over a time period corresponding to reproductive events (e.g., using data obtained from post-breeding surveys or censuses, which for example might occur annually), yet EPA appears to want to model daily population changes even for long-lived, annually reproducing species such as some fish. Daily incrementing will be necessary to capture the rapid changes in survival during exposure to time-varying concentrations of contaminants. For continuously reproducing species, daily incrementing of reproduction at a constant rate might be appropriate. However, this approach will produce a misleadingly optimistic estimate of recovery time for species that breed only periodically, especially annual breeders.
2. Age of first reproduction can be altered by exposure to pollutants. Changes in age of first reproduction can easily be incorporated into age-based population models but are not as easily incorporated into stage-based population models. I could not see a proposal to address this important concern.

Until appropriate and extensive enough datasets are available to incorporate density dependence into population modeling of the effects of contaminants, I suggest the simpler assumption of exponential growth be adopted. I believe EPA should “walk” until it is equipped well enough to “run”. Assuming exponential population growth, the modeling task becomes considerably easier. The simplest approach to advancing the criteria methods would be to assume constant-exposure concentrations (as is done in the current criteria method) and calculate the population growth factor (λ) from data on survival and reproduction that are, for example, already collected in many chronic toxicity tests with invertebrates and fish. No time-series modeling would be necessary, and methods exist for estimating the uncertainty associated with λ (e.g., Meyer et al. 1986, 1987, Caswell 1989). The next more complicated approach might be to try to incorporate the effects of time-varying exposure to contaminants into the calculations; however, as I stated above, it is not clear how reproduction is affected by time-varying exposures. This would require time-series modeling, and I don’t believe even that type of density-independent modeling is feasible now (because of a lack of appropriate data and appropriate understanding of how to model changes in reproduction during time-varying exposures).

As a major drawback to the exponential-growth approach, the authors of the Water-Based Criteria document asserted that a recovery time cannot be defined if exponential population growth is assumed. Although by their definition of recovery, the authors are correct that an impacted population can never match the control population size measured

at the same time, I disagree that a useful index of recovery time cannot be calculated. The first ad hoc approach that comes to my mind for defining recovery time would be to ask: How many days (or years) would it take a contaminant-impacted population to reach the size that the control population had at the time the exposure to the contaminant is removed from the exposed population? For example, assume exposure to the contaminant over a several-day period causes the size of the impacted population predicted by an exponential-growth model to decline to 100 individuals while the model-predicted size of the unexposed population increases to 1000 individuals. How many days after the contaminant exposure ends would it take the previously exposed population to increase from 100 to 1,000 individuals? That might be a useful index of recovery time. I'm sure other ways could be devised to measure recovery time, too.

Charge Question 2.3: Please comment on whether the proposal for aggregating effects across species being considered by EPA is scientifically appropriate for use in deriving water-based criteria.

I believe the proposal to aggregate effects across species through what amounts to an averaging of time-weighted average percentage impairments is innovative but is only one of several possible approaches, all of which have advantages and disadvantages. I don't think it is scientifically inappropriate.

Although EPA is trying to convince us that this aggregation approach circumvents problems associated with species sensitivity distributions (SSDs), the aggregation approach seems to carry some of the same baggage associated with SSDs. For example, if the assemblage of species included in an aggregation calculation is not representative of the assemblage EPA intends to protect, the outcome will be biased. Because a relatively small number of species will have adequate data for modeling the growth of their populations, it would be relatively easy to bias the aggregation outcome by funding appropriate studies on one or a few very sensitive or insensitive species (analogous to what can be done to bias SSDs, but much easier because of the smaller number of species that probably will be in the proposed assemblages).

Charge Question 2.4: Please comment on whether the framework being considered by EPA for deriving water-based criteria is scientifically appropriate for use in deriving the criteria.

I am pleased that EPA is attempting to incorporate more geochemical, physiological, and ecological principles and understanding into their aquatic life criteria. Although I do not believe the full extent of the proposed changes is feasible at this time, I do not want to discourage EPA from striving to achieve their goals. In general, the approach is scientifically appropriate. However, I suggest that EPA assess the feasibility of developing credible procedures and datasets for these lofty goals within a reasonable time frame; and if the goals do not appear to be achievable, EPA instead should re-gear to enact workable, scientifically credible, but incremental improvements to the criteria methods. Otherwise, the 1980s approach to criteria might by default remain with us for a long time.

Unless considerable effort and funding (within EPA and/or industry) is devoted to achieve these lofty goals, I fear the criteria revisions will remain a twinkle if EPA's eyes. I recommend a major thrust within EPA to demonstrate proof of principle for the proposed methods with one or a few priority pollutants (i.e., demonstrate that the proposed approach is workable). Based on that effort, EPA could then make appropriate modifications to the proposed criteria methods and estimate the amount of money and personnel needed to develop appropriate datasets for all chemicals for which revised criteria are needed, at each of several possible levels of criteria improvement.

A major drawback of the proposed approach is the omission of consideration of non-equilibrium partitioning of contaminants in the water column (and perhaps in sediment and food). Non-equilibrium partitioning can be especially important for metals, because contact times greater than or equal to 24 hours are necessary to equilibrate free metal ions (e.g., Cu^{2+}) with dissolved organic matter (DOM) under realistic conditions in some natural waters (Ma et al. 1999). Therefore, concentrations of free metal ion in the water column after only a few hours of contact time with DOM can be at least an order of magnitude higher than the concentrations of the same free metal ion after 24 hours of contact time. I suspect a wide spectrum of metal-DOM contact times can be found in the real world. Therefore, before most of the knotty subtleties of time-varying exposures to contaminants and density-dependent population growth are addressed, I believe the very important reality of non-equilibrium partitioning of metals (and perhaps some organic compounds) should be addressed and incorporated into aquatic life criteria. If we can't get the geochemistry approximately right, we most assuredly run the risk of having greatly under-protective and/or greatly over-protective criteria.

Literature Cited

- Caswell, H. 1989. *Matrix Population Models*. Sinauer Associates, Inc., Sunderland, Massachusetts, USA.
- Ma, H., S.D. Kim, D.K. Cha and H.E. Allen. 1999. Effect of kinetics of complexation by humic acid on toxicity of copper to *Ceriodaphnia dubia*. *Environmental Toxicology and Chemistry* 18:828-837.
- Meyer, J.S., C.G. Ingersoll, L.L. McDonald and M.S. Boyce. 1986. Estimating uncertainty in population growth rates: Jackknife vs. Bootstrap techniques. *Ecology* 67:1156-1166.
- Meyer, J.S., C.G. Ingersoll and L.L. McDonald. 1987. Sensitivity analysis of population growth rates estimated from cladoceran chronic toxicity tests. *Environmental Toxicology and Chemistry* 6:115-126.
- Schnoor, J.L., J.P. Connolly, D.M. Di Toro, N. de Rooij, M. Diamond, R.S. Kinerson Jr., D.B. Porcella, W.L. Richardson and J.F. Stine. 1997. Environmental fate and transport. pp. 71-80 in: H.L. Bergman and E.J. Dorward-King (eds.). *Reassessment of Metals Criteria for Aquatic Life Protection*. SETAC Press, Pensacola, Florida, USA.

Initial Comments for SAB consultation on Proposed Revisions to Aquatic Life Guidelines

Judy L. Meyer

9/19/05

Overview of Proposed Revisions to the Guidelines for Deriving Ambient Water Quality Criteria for the Protection of Aquatic and Aquatic-dependent Wildlife

Charge question 1.1 Use of risk assessment framework

The task of revising the guidelines is an important one, and the committee is headed in the right direction. I am particularly pleased to see the recognition of the potential importance of dietary uptake in these revisions. Moving beyond total reliance on water-based criteria is a much-needed improvement in environmental protection.

I am concerned that there is no timeline in any of the documents that I have read. It has been two decades since the guidelines were first released, this committee has been working intermittently for at least 10 years, and yet the documents contain many statements of problems or questions that the committee has not discussed or has not achieved a resolution. I am very worried that committee has taken on a very large task, and by trying to complete everything at once, it will be another several decades before any changes are made. That would be disastrous. A staged or incremental approach would seem more appropriate for revisions of the magnitude that are being considered. Constructing a timeline might help identify components that could be accomplished with a stepwise approach.

The approach described is based on toxicity testing with little attention being given to using reference conditions for the region to set criteria. The bioassessments being done as part of the Tiered Aquatic Life Use (TALU) could be helpful in identifying assemblages that are not impacted. Establishing ranges of contaminant concentrations at these sites and in a variety of species at these sites would provide a basis for establishing criteria that are protective of aquatic life. I am not saying that the toxicological approach should be abandoned (!), but I do think use of reference conditions offers an alternative pathway that should be considered. How well is the work of this committee integrated with what the TALU program is doing? Use of information like this could be very useful in identifying impaired reaches and determining if contaminants are a likely cause of impairment. I discuss this further below and under tissue-based criteria.

There are several things that are not included in the framework (Figure 2) that are worthy of consideration:

(1) How can the process described with this framework be better integrated with the TALU bioassessment work that EPA is doing? It seems that information being generated using bioassessments in this program through sampling of biotic assemblages at sites would be extremely useful in the problem formulation and effects assessment steps. The nature of changes in the fauna observed at sites provides necessary insight for problem formulation. The approach being described is toxicological; some ecological balance is needed to make the criteria protective of the organisms actually found in aquatic ecosystems being protected. Better integration of this process with findings from TALU bioassessments would provide some of that needed balance.

(2) Little mention is made in any of the documents (or in the figure) on representativeness of the species selected for testing and upon which the water quality criteria will be based. What criteria are to be used to decide what a species represents and what attributes need representation? The lists given include the broadest of taxonomic representation (invertebrate, vertebrate, plant), some mention of trophic position (herbivore, omnivore, piscivore), and perhaps habitat (benthic vs. planktonic). These seem like extremely broad categories. For example, I am concerned that there are little data from the more sensitive benthic insects from streams, and that feeding mode (e.g. filtering collector, collector-gatherer) has not been considered. Part of problem formulation should be what are the species characteristics (maybe it is lipid content, size, feeding mode, reproductive habits, etc.) that need to be represented in the mix of species used to establish the water quality criteria. I am not convinced that sophisticated statistical procedures will make up for a truly representative set of species used in the analyses. Once again, reference to some of the biotic sampling being done for TALU could be helpful here.

(3) One of the management objectives includes the following endpoints: "sustainability, ..., function, productivity, stability ... of aquatic communities and ecosystems." This list needs to be considered when deciding on endpoints. It is not clear to me how these more functional assessments will be incorporated into the guidelines.

(4) Problem formulation would be a good place to consider the mixture of compounds experienced by the biota at a site. A recognition that the biota are experiencing mixtures of contaminants would be part of an hypothesis "about why ecological effects have occurred." The questions in the first box in Table 1 address this. To those questions I would add: What other pathways of action impact this endpoint? Is the chemical likely to occur with other chemicals with opposing mechanisms of action?

(5) It is not entirely clear to me where food web effects are in Figure 2. Is that meant to be encompassed in the community and ecosystem models that translate risks to populations into the risks to ecosystems?

Charge Question 1.2: Use of three types of criteria

This approach was needed yesterday! There is abundant scientific evidence that there are multiple pathways by which contaminants impact biota, and that criteria based exclusively on contaminant concentrations in water are not protective. So I applaud this proposed direction of change in the guidelines. I'll comment on each criterion:

Water-based criteria

I am concerned that meeting the increased data needs for the kinetic model will come at the expense of attention to taxonomic diversity. Taking this approach may result in very detailed understanding of toxicity in a couple species, yet little understanding of the diversity of responses to be expected in a real stream or lake. The amount of work still needed to begin to implement this approach carries with it the danger that it will be a long time before we see new guidelines. Once again, I recommend consideration of a stepwise presentation and implementation of guidelines.

The description of the TEA model once again brings up my concern about representativeness of the suite of surrogate species used in the assessment. What do the

surrogates represent: reproductive mode, lipid content, size, feeding mode, etc.? Much more attention needs to be paid to choice of species to be used in the assessments.

I can understand how the kinetic toxicity model could improve the ability to establish criteria. What was not clearly explained was how this model relates to enforcement of the criteria. Exceedence criteria would need to be more stringent for a contaminant that was measured monthly than for one that was measured daily.

I didn't see much discussion of estimating uncertainty. As models grow more complex with more parameters, uncertainty is likely to increase.

Tissue-based criteria

Tissue-based criteria could be used in two ways; yet the committee is exploring only one of those. The committee is relating tissue concentrations to toxicological endpoints. Another way to use tissue-based criteria is in identifying sites where particular contaminants may be a problem (e.g. in the initial problem formulation step). Using tissue-based criteria in this manner would involve characterizing tissue (or whole body) concentrations at reference (uncontaminated sites) and comparing observed concentrations at other sites to the range of concentrations measured at the reference sites. This, coupled with evidence from bioassessments (which would suggest types of species that are missing) should provide useful information on which contaminants are of concern at a site.

Contaminants that do not biomagnify but are bioaccumulated can still have an impact on higher trophic levels. For example, the study by Woodward et al. (1995. Canadian Journal of Fisheries and Aquatic Sciences 52: 1994-2004) showed that young trout that fed on benthic invertebrates that had bioaccumulated Cd had lower growth rates than trout fed on non-contaminated benthic invertebrates. Findings such as these did not seem to be incorporated into the documents that I have read.

The biodynamic model proposed by Luoma and Rainbow (2005. Environmental Science and Technology 39: 1921-1931) seems to offer some promise for predicting bioaccumulation and its effects on organisms. I suggest the committee consider incorporating this model into their discussions.

Taxon-specific criteria

Use of taxon-specific criteria represents a needed advance. I presume this kind of analysis could be used for species that are identified as keystone species or habitat-creating species.

Greater attention needs to be paid to identifying protocols for determining appropriate surrogate species.

In the problem formulation step, another factor that needs to be included in the effects analysis is any effect of the contaminant on a food resource of the species of concern.

L.S. McCarty Scientific Research & Consulting 1156868 Ontario Inc.
94 Oakhaven Drive, Markham, Ontario, Canada L6C 1X8 Phone 905/887-0772 lsmccarty@rogers.com

September 19, 2005

Thomas Armitage, Ph.D.
Designated Federal Officer
USEPA Science Advisory Board (1400F)
1200 Pennsylvania Ave., N.W.,
Washington, D.C.
USA 20460

Dear Dr. Armitage:

Please consider this letter my draft response to the charge questions concerning Tissue-Based Criteria for “Bioaccumulative” Chemicals for the Aquatic Life Criteria Guidelines Consultative Panel. I am still working on some of my responses to the other charge questions so I haven't included them here. However, as the response to the framework scope is in reasonable shape, I have also included it. My final response to all charge questions, including a final version of these comments will be provided later.

Charge Questions

1. **1. Scope of the Proposed Framework for Revising Aquatic Life Water Quality Criteria**
2. 1.1 Please comment on the use of the Guidelines for Ecological Risk Assessment as an essential and relevant organizing framework for development of science-based criteria for the protection of aquatic life and aquatic-dependent wildlife. Does the SAB have any specific recommendations on how to improve or clarify the generic conceptual framework diagram?

The material on this issue appears in the draft document entitled “Overview of Proposed Revisions to the Guidelines for Deriving Ambient Water Quality Criteria for the Protection of Aquatic and Aquatic-Dependent Wildlife”. Although I am in general agreement that the USEPA ERA Guidelines should be used in developing water quality criteria, I think that a modest revision could improve and clarify the process enabling, among other things, better risk communication.

The first issue is Figure 1 on page 9 where an aquatic life criteria development paradigm is presented. I continue to be uncertain why the U.S. EPA does not include risk management goals and objectives as an integral component of the ERA paradigm. Especially as it is clear from the text in Section 2.3 on page 11 that management goals and objectives determine both the overall nature and details of the risk assessment process and risk management actions. I have commented on this issue previously (McCarty and Power, 1997; Power and McCarty, 1998) and suggest the same revision herein. Goals and objectives are a key component of any risk paradigm as the overall process is risk management. Risk assessment is a component of and informs the risk

management process only to the nature and extent indicated in the risk management goals and objectives. Figure 1 below is a suggested revision of Figure 1 on page 9.

This revision clarifies the relationship between the risk management and risk assessment portions of the paradigm, especially that the overall result of the process is to select an action to achieve the stated risk management goals and objectives. I have changed the arrow between risk characterization and exposure assessment to a double-headed arrow with a dashed line which reflects exposure issues discussed in the draft guideline document. The two-way information flow here indicates that iterations between the lower three boxes (exposure, characterization, actions) are possible and probably likely.

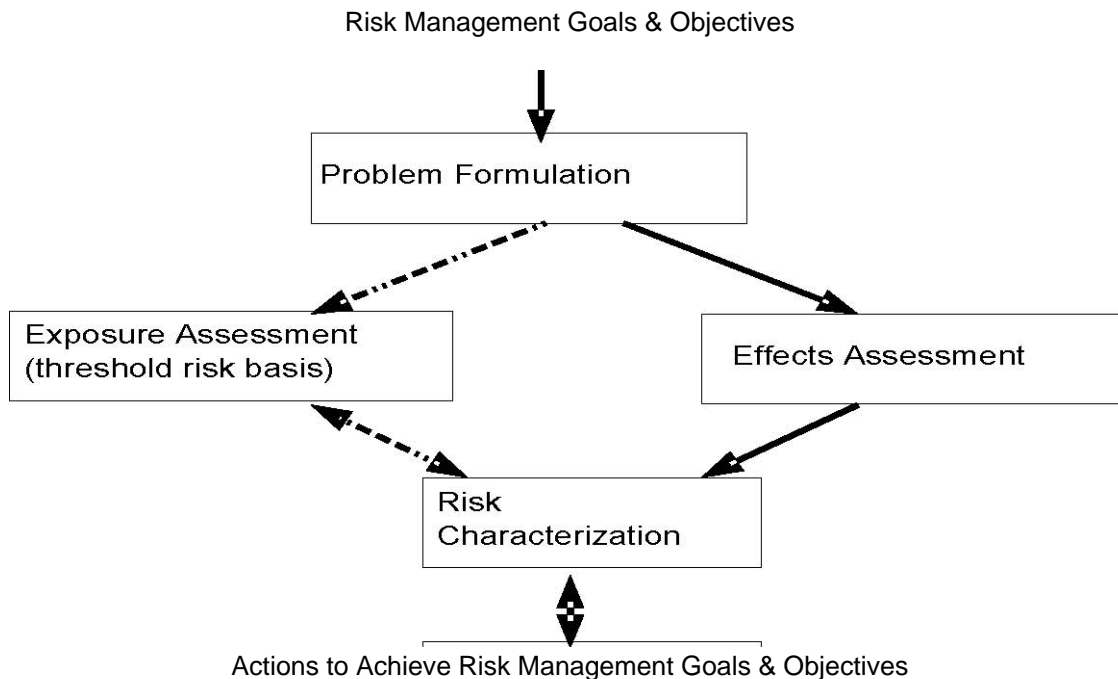


Figure 1: Suggested Revision of USEPA Aquatic Life Criteria Development Paradigm

The second issue is related to Figure 2 on page 17. This is the conceptual framework diagram for chemical criteria derivation and risk assessment. Although it contains the appropriate information for this exercise, and I recognize that it is trying to present a summary of the various processes involved that justifies the three types of criteria development suggested, I find it confusing. I have created a revised version of this conceptual model which is presented in Figure 2.

The revised conceptual model presented in Figure 2 does not contain all of the detail in the original. With further work such information could be added; however, the intent of the revision at this stage is not to be entirely prescriptive. Rather changes are made to clarify the basic relationships presented in the original and fundamental to understanding toxicology and its use in regulatory activities. The key components of interest to EPA – media-based models, tissue-based models, and taxon-specific differences – remain. The main rationalization is in the various media-based models. Rather than separation, as in the original, they are combined. This is because all media-specific models describe a similar process of exposure, partitioning, and response to an accumulated internal exposure.

Although tissue-based models explicitly use some measure of accumulated exposure internal to the organism, media-based models, whether they consider exposure water (respiratory and skin exposure), ingestion (diet, water, sediment) or both, implicitly consider internal exposure in relating exposure to effects. Water-and tissue-based approaches are much more similar than they are different. Although separate water-based and tissue-based approaches to criteria development is an acceptable limitation for the purposes of regulatory development, the basic nature of the relationship between received internal dose and the induced adverse effect is essentially identical. In fact, increased knowledge about the received dose-effect relationship would be of significant value, irrespective of whether the dose metric is media-or tissue-based, and such knowledge may also be transferable between the two.

The specific separation of toxicity evaluation in a species by species manner reflects the actual nature of the process and highlights the importance of the taxon-specific approach. However, is using at species specific approach extreme caution must be exercised in ensuring that real differences or similarities in species toxicity is concluded only after adjustment of the experimental data for differences caused by physical, chemical, and biological factors interacting with differences in the specific toxicity test protocols employed for the different species.

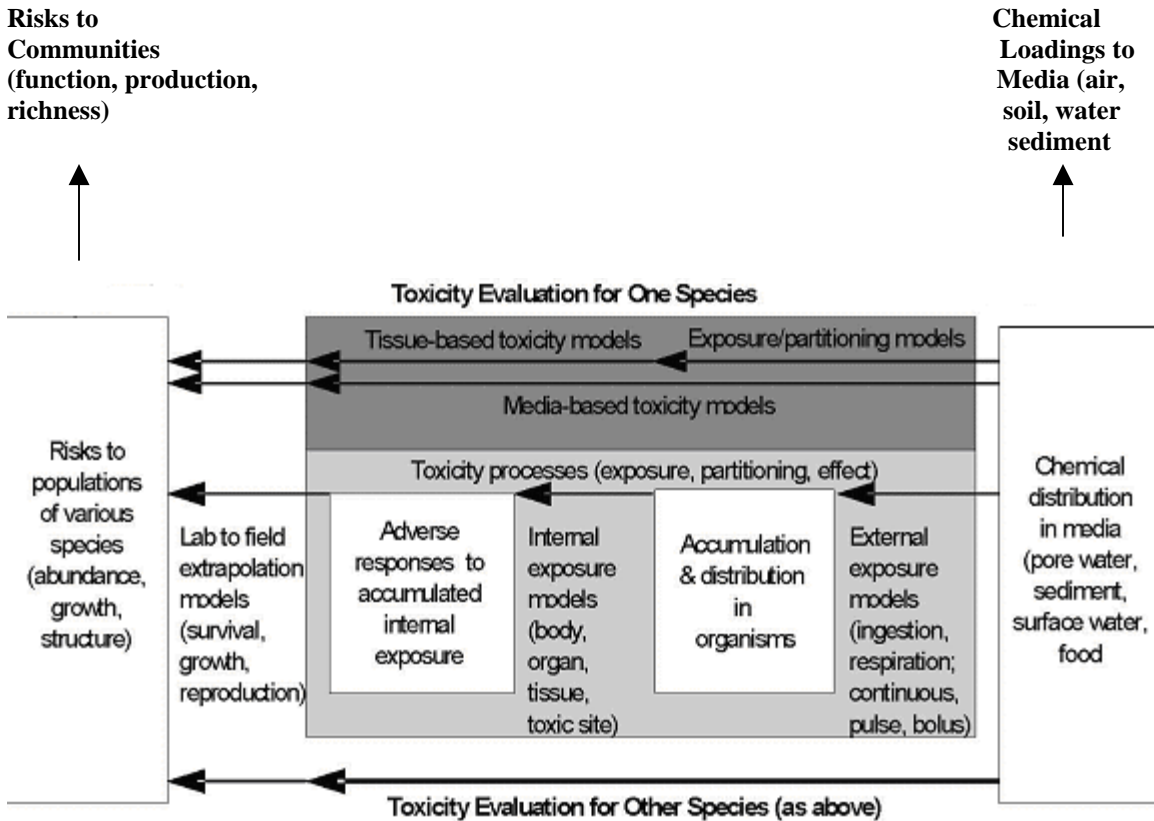


Figure 2: Revised Conceptual Model for Chemical Criteria derivation and Risk Assessment

- 1.2. Please comment on whether the proposed criteria types and the scientific focus for each criteria type are logical and scientifically valid for developing a holistic and integrated criteria framework.

The use of three types of criteria – media-, tissue-, and taxon-based – in deriving water quality criteria does have some logic to it, although I believe that the taxon-specific issue will ultimately become a subcomponent of the other two criteria types. Also, as noted in my response to Question 1.1, the water-and tissue-based approaches are essentially the same, with the latter directly and the former indirectly using an organism-based dose metric. However, my main difficulty is not with the general scientific focus. Although the bulk of the issues and concerns have received extensive consideration, there is a fundamental issue that is not receiving sufficient focus; specifically, the generally poor level of toxicity test data interpretation. As indicated the revised figure presented in Figure 1, the toxicity test interpretation under discussion is separate from laboratory to field extrapolation issues, although, since laboratory to field extrapolation depends on thorough toxicity test interpretation, it will be affected by any interpretation difficulties.

It is important to recognize that virtually all regulatory development schemes, and much related scientific literature, employs (usually implicitly) the assumption that the various toxicity test data employed and manipulated in various processes are compatible and readily comparable in nature. Expressed in another way, it is effectively assumed that experimental toxicity test data from generally-accepted testing protocols are the toxicological equivalent of physical-chemical properties of substances. That is to say that it is largely a constant, albeit with some measurement variability. This assumption is not valid (Mackay et al., 2001).

Briefly, some properties of chemicals, termed intensive, depend on the nature of the substance and not its quantity. Examples include boiling point, solubility in water, and octanol/water partition coefficient (Kow). Other properties are termed extensive as they depend on the amount of chemical present. Concentration is an extensive property as it varies with the quantity of substance present. The induction of a toxic effect in an organism depends on the chemical nature of the substance, the nature of the organism, and the dose. While the nature of the substance and organism can be considered intensive or quasi-intensive dose is extensive since it depends on amount (i.e., concentration).

Thus, toxicity as measured in standard testing is a function of both intensive and extensive quantities and cannot be intensive. However, it is possible to create a quasi-intensive property by dividing one extensive property by another to ratio out the extensive amount. Toxicologists employ such an approach using the ratio (extent of occurrence of a defined toxic event) / (quantity or concentration of chemical). Thus, toxicity metrics such as LC50 are most accurately termed measures of specific toxicity, which are quasi-intensive. Quasi-intensive properties such as specific toxicity can only be considered intensive under specific reference conditions. When determined under different sets of reference conditions they are not directly comparable with each other.

This issue is of direct relevance in ranking relative toxicity or risk of substances and other regulatory activities since experimental or calculated toxicity test data are key information in such processes. To be valid comparisons of relative potency must be made with data collected with exactly the same toxicity testing protocol. Among other things the age/size of organism tested, the duration of exposure, and adverse effect endpoint should be the same or very similar for data to be considered comparable measures of specific toxicity. It is likely that as differences between testing protocols become greater, the data become less comparable and greater error is present in the analysis. In other words, different toxicity test protocols produce different measures of

toxicity due to differing physical, chemical, biological, and/or temporal modifying factors. Thus, quasi-intensive toxicity data are comparable within a specific toxicity testing protocol (i.e., identical except for exposure to different substances). However, to combine and jointly analyze data from several different toxicity testing protocols, the toxicity test results must be converted to a common reference condition, which corrects for the differing influences of various modifying factors in the various test protocols..

An analogy may be helpful. Consider the measurement of length. There are many units both Metric (metres, centimetres, etc), Imperial (yards, feet, inches, etc), plus various traditional units (cubits, paces, etc.). Given a list of measurements taken with a variety of length units it is impossible, without converting the diverse units to a common basis, to determine if a data set contains repeated measurements of an identical length or separate measurements of a variety of lengths. For precise conversions the measurements must also be adjusted to a common temperature as materials change length with different temperatures. Unconverted lists containing diverse units could have the differences attributed to uncertainty, variability, or measurement errors, especially when the units are not dramatically different, e.g., metres and yards, inches and centimetres. However, in the end, without conversion, mistakes will be inevitable due to either real data trends being obscured or false ones arising.

There is a specific example related to a scientifically sound technical approach for examining relative species sensitivity and species tolerance distributions in aquatic risk assessment.

“Regarding the quality of toxicity data, one has to keep in mind that artificial differences in sensitivity between species may result from the use of a standardized, arbitrary exposure time, the indiscriminate use of different effect parameters (growth, reproduction, survival), and the ignorance of sensitive life stages.

The last drawback is that there is no theoretical basis behind the proposed probability distributions of the assessment factors, as they are selected empirically. Particularly for the AFinterspecies, the selected distribution may be questionable.” (Roelefs et al. 2003, pg. 1392)

That is to say, summarizing the above statement, when using uncorrected or unadjusted standard toxicity test data from a variety of test protocols, it may well be that much of the variability which is often attributed to differences in species sensitivity, is largely an artifact of the differences in the testing protocols and the lack of correction for those differences. A similar criticism can be made regarding multi-species, multi-protocol frequency distribution analysis of acute to chronic toxicity ratios. Currently, much of the variability, uncertainty, or “noise” in available toxicity data is addressed in regulatory development guidance by policies which include schemes incorporating application, uncertainty, safety, assessment, and/or extrapolation factors. Such schemes rate quality/quantity of toxicity data, employ various acute to chronic extrapolations, and evaluate differences in toxicity test results within and between various test species then apply large (orders of magnitude) factors to a selected toxicity estimate to generate regulatory numerical guidance in the form of media (water) concentrations. The issue of lack of compatibility without correction for test-protocol influences affects all aspects of regulatory criteria development since ranking relative toxicity is the key activity in any criteria development process. However, as noted above, a substantial portion of the variability, uncertainty, or “noise” in available toxicity data is likely related to the lack of correction for test-protocol-induced differences. It follows that, if the data used in such processes is corrected for much of the inter-protocol-induced variability, substantial refinement in regulatory protocols should be possible enabling the use of smaller and more precise application/assessment factor schemes.

Although this issue appears daunting, there are approaches to address it. The essence of the challenge is to account test-specific differences in physical, chemical, temporal, and biological factors such that the corrected data primarily reflect differences due to different chemical substances tested and different species sensitivity. Although it is unlikely that this can be done in any absolute sense or for all types of toxicity tests, some substantial improvements should be possible. Three areas related to better interpretation and analysis of experimental toxicity test data come immediately to mind:

- time-toxicity analysis
- avoidance of response endpoint lumping
- development of a species-specific mode of action classification scheme

The establishment of time-independent or time-invariable toxicity estimates via time-toxicity curve analysis within a testing protocol has been long advocated. For example, Sprague (1969) recommended that toxicity tests continue until a threshold or incipient value is achieved. Although such values can often be estimated by inspection of probit-time curves, curve fitting programs are available (see Rand et al., 1995, pg 33-34). In addition to providing threshold toxicity estimates analysis of time-toxicity curves provides a key one compartment, first-order kinetics (1CFOK) model parameter, the overall elimination rate (k_2). Further exploitation of this information allows better determination of effective exposure durations and provides a means begin to evaluate the influence of toxicological modifying factors such as body size and metabolic degradation (Rand et al., 1995; Mackay et al., 1992). However, the estimates are only time-independent within the context of the testing protocol in question as this is a measure to ensure that a steady-state has occurred between the exposure concentration in the medium and the received dose in the exposed organisms.

Adverse effect endpoints are typically grouped into three categories: mortality/survival, growth, reproduction. These endpoints groupings are useful in that they are key factors in population dynamics and modelling populations and are vital aspects in any laboratory to field estimation process. Although there is often extrapolation between the categories (i.e., acute to chronic ratios) the foundation for this is uncertain. Mayer et al. (1994 and other papers) provided an explanation for the presence of the acute to chronic ratio phenomena where the same mode of toxic action is causing both the acute and chronic effects. They found that the chronic endpoints used in early life stage fish tests were similar to the lower tail of acute toxicity tests (i.e., LC0.01) conducted with the same species and chemical. The relationship was valid for lethality and probably growth but not for reproduction. In fact, they commented:

“Finally, chronic toxicity tests are not necessary if lethality, and probably growth, are the only end pointsof interest in fishes – predictions from well-conducted acute tests are very adequate” (Mayer et al., 1994)

Similar work has not been carried out with other organisms. Although it is likely that similar relationships exist between acute and chronic lethality for other organisms, this has not yet been determined. Also, for some other common test species (Daphnids, rotifers) chronic tests typically use growth and/or reproduction endpoints where the acute-chronic relationship was not valid in fish investigations. All of this suggests that additional care must be taken with acute to chronic extrapolation. In particular, in regulatory criteria development guidance available toxicity testing protocols should be clearly classified into one of the three groups -mortality/survival, growth, reproduction – and separate extrapolation procedures prepared.

In addition to variable effects on the expression of toxicity due to chemical characteristics that control accumulation kinetics (toxicokinetics), differences in the expression of toxicity are also contributed by different mechanisms/modes of toxic action (toxicodynamics) (e.g. McCarty and Mackay, 1993). Although a number of attempts have been made no generally-accepted, widely applicable mode of toxic action scheme is currently available (McCarty, 2002; Borgert et al., 2004). A key action is to establish a simple mode of toxic action (MOA) scheme and evaluate the influences of MOA on toxicity test results, both kinetics and dynamics. This is a critical to enable improvements in toxicity data interpretation and regulatory development guidance.

It is unrealistic to expect full consideration of MOA in the near future due to the current poor state of the art. However, it should be possible to use existing information and simple preliminary MOA schemes to provide some preliminary guidance. This would entail building enhanced 1 compartment or 2 compartment exposure-based toxicokinetic models (EBTK) models for the several common test species such as fathead minnow, Daphnia etc.. Then using the best available toxicity data, and a thorough analysis of that data, establish “expected” results for several chemicals which are characteristic of categories in a simple proposed MOA scheme. Nonpolar narcosis is the baseline for organic chemical toxicity and agreement on few specifically-acting organic toxicant categories should be possible. Inorganics and metals will require similar consideration and classification. Once the models are established examples of the influence of some common modifying factors such as body size, metabolism, and temperature can be demonstrated and checked against available data.

MOA influences typically employed evaluation approaches such as acute to chronic ratios, time to steady state conditions, and differences in species sensitivity. Ultimately in any MOA scheme the likelihood that some test species do not exhibit a certain MOA must be addressed. For example, some herbicides are specifically toxic to certain plants but cause toxicity to animals by another MOA such as narcosis. Similarly, some specifically-acting pesticides are unlikely to cause toxicity by the same MOA to all animals, or at least there may be a substantial difference in the received dose potency. Although any MOA scheme is likely to undergo many iterations in its development, it is important to organize current knowledge and data in this manner to determine trends and relationships and identify knowledge and data gaps. As well as providing a path forward for scientific hypothesis formulation and testing, it would provide considerable guidance to those charged with development of regulatory criteria.

Although the above brief review cannot do justice to the real nature and magnitude of the toxicity test interpretation issue, it does suggest how the problem might be attacked and some of the benefits that might accrue.

3. Tissue-Based Criteria for “Bioaccumulative” Chemicals

- 3.1. Please comment on the rationale and conceptual approach used for the development of tissue-based criteria for this group of chemicals. Is the SAB aware of other approaches for deriving criteria for these bioaccumulative chemicals that EPA should consider?

I have long supported the regulatory approach advocated by Don Mackay. That is to determine the fate and distribution of chemicals in various compartments of the environment, develop an understanding of the adverse effect of interest in the target environmental compartment (the organism), then use knowledge on the inter- and intracompartments fate and distribution to determine the best environmental compartment for promulgating regulatory guidance.

Traditionally water concentrations have been the regulatory compartment of choice for aquatic systems, but the discovery of very hydrophobic contaminants such as 2,3,7,8-TCDD highlighted the difficulty of analytical detection of such chemicals in water. On the other hand, much higher levels of 2,3,7,8-TCDD in the bodies/tissues of exposed organisms made estimation of levels in organisms more reliable. Thus, organism concentrations became a more useful means to express regulatory guidance for the most affected organisms. However, water levels of 2,3,7,8-TCDD are an important aspect in the overall aquatic foodchain accumulation process and the entire foodchain accumulation process from water concentrations through water and dietary exposure foodchain, must be understood to reliably estimate levels associated with adverse effects. Given this understanding the regulated medium could still be water, but the analytical challenges still mitigate against it.

Simply put, to provide sound scientific input into environmental regulations the interrelationships between distribution, fate, exposure, and effects of chemicals must be understood in some detail. If these interrelationships are understood in the required detail any compartment -soil, air, water, sediment, organism -can be used for regulatory purposes. The choice is largely a function of cost and convenience. For the case of 2,3,7,8-TCDD it makes more sense to measure organism tissue levels than to try to routinely measure the very low water concentrations which require expensive sampling and analytical protocols. Although both water and tissue chemical analysis equipment and techniques for TCDD determination have developed substantially in last quarter century since TCDD became an issue, the principle still holds.

The short discussion above is simply to note that shifting from a water-based dose surrogate to a organism/tissue-based surrogate is not a panacea, nor is it really a dramatic change. Much has to be known about chemical fate and distribution in the environment and in organisms, effects of accumulated concentrations in exposed organisms, and extrapolation of controlled laboratory data on various single species to populations and communities in the real world. In fact, the knowledge requires should be sufficient to provide regulatory guidance based on either water or organism levels. Thus, the key philosophical challenges faced in preparing tissue-based criteria for “bioaccumulative” chemicals are exactly the same as faced by those preparing water-based criteria. As pointed out in the draft US EPA guideline revision document for “bioaccumulative” substances “Note: the tissue or water-based approaches are not mutually exclusive, ...” (p. 12)

For the sake of consistency it is important that both water-based and tissue-based approaches use comparable and compatible philosophical and technical processes and procedures for evaluating, analyzing, and modelling fate, distribution, and toxicity. Although the draft documentation suggests that such coordination is an important goal, it will be a challenge to ensure that this is carried through when the detailed technical procedures for developing specific aquatic life criteria are formalized.

Overall, the rationale and conceptual approach discussed in the draft US EPA document largely reflects the current state of knowledge and critical issues related to criteria development for “bioaccumulative” substances in the aquatic environment. However, being aware of the data, knowledge, and critical issues is not a guarantee that the ultimate recommended criteria development protocol will properly or adequately employ that data and knowledge or adequately address the critical issues. Expert peer review and guidance will be needed throughout the preparation of the derivation guidance to ensure the technical aspects of best employ existing data and knowledge and optimally address the critical issues.

- 3.2 Considering the strengths and limitations of the more flexible approach used to derive tissue-based criteria, please comment on the rationale and preference for allowing

flexibility in the procedures used?

The flexibility suggested, that of being able to use deterministic approaches to characterize toxicity data when data are limited and probabilistic approaches (e.g., species sensitivity distributions) when data are sufficient, appears to be reasonable. Additional guidance for site-specific refinement will be a useful enhancement. However, a key concern regarding flexibility and the use of different approaches for different substances based on data quantity/quality, is the overall consistency and compatibility of the results of the criteria development process. That is to say, what are the differences expected from the use of different protocols for different chemicals: is the level of protection similar or different?

It is not possible to comment on this in any detail at this time as, although a number of important issues have been raised, specific approaches have not been determined in some cases. For example, as discussed in Section 3.2.2 and 3.2.4, a technical basis for addressing differences in species sensitivity has not been addressed for the deterministic approach. For the probabilistic approach the species-specific distribution (SSD) method is contemplated, as is the toxic effect aggregation (TEA) method, but which method or the details of applying a chosen method are not presented.

In summary, although I am generally in favour of flexibility, it is not possible to evaluate the specific influence that flexibility might have in this case as the alternative methods are not clearly established.

3.3. Please comment on the rationale used by EPA for determining if/when to use population modelling in the development of Tissue-Based Criteria?

I agree that the substantial additional effort required to carry out population modelling is currently not warranted for the case of national-level tissue-based criteria as it is unlikely to add appreciably to the risk estimation process. The discussion in section 4.6 Role of Population Modelling presents the issue clearly and succinctly. A recent review on the issue of modelling stress in fish populations provides a more detailed examination of the issue of population modelling in environmental risk assessment, and is equally applicable to the water-based criteria approach (Power, 2002). Population modelling should be a key element in environmental regulation methodology. Unfortunately the state of the art is such that it requires considerable additional effort, yet provides only questionable and uncertain changes to the outcome such that is difficult to determine if any improvement is obtained. .

I agree that there is a role for population modelling in site-specific criteria development. Further development of population modelling in this context, and development of general guidance for its use in site-specific situations, appears to be a reasonable approach. The opportunity to enable the development of additional case studies should aid in the much-needed refinement of the modelling of population and community responses to stress.

Literature Referenced

Borgert, CJ, TF Quill, LS McCarty, and AM Mason, 2004. Can mode of action predict mixtures toxicity for risk assessment? *Toxicol Appl. Pharmacol* 2001:85-96.

Mackay, D, LS McCarty, and M MacLeod, 2001. On the validity of classifying chemicals for persistence, bioaccumulation, toxicity and potential for long range transport. *Environ Toxicol Chem* 20(7)1491-2001.

Mackay D, H Puig, and LS McCarty, 1992. An equation describing the time course and variability in uptake and toxicity of narcotic chemicals to fish. *Environ Toxicol Chem* 11:941-951.

Mayer, F.L., G.F. Krause, D.R. Buckler, M.R. Ellersieck, and G.H. Lee, 1994. Predicting chronic lethality of chemicals to fishes from acute toxicity test data: Concepts and linear regression analysis. *Environ. Toxicol. Chem.* 13:671-678.

McCarty, LS and M Power, 2000. Approaches to developing risk management objectives: an analysis of international strategies. *Environ Sci Policy* 6:311-319.

McCarty, LS and M Power, 1997. Environmental risk assessment within a decision-making framework. *Environ Toxicol Chem* 16(2):122-123.

Power, M and LS McCarty, in press. Environmental risk management decision-making in a societal context. *Human Ecol Risk Assess.*

Power, M and LS McCarty, 1998. A comparative analysis of environmental risk assessment/risk management frameworks. *Environ Sci Technol* 32(9):221A-231A.

Rand, PG Wells, and LS McCarty 1995. Chapter 1: Introduction to Aquatic Toxicology. IN *Fundamentals of Aquatic Toxicology II*, GM Rand (ed). Taylor and Francis, Bristol PA.

Roelofs, W., M.A.J. Huijbregts, T. Jager and A.M.J. Ragas, 2003. Prediction of ecological no-effect concentrations for initial risk assessment: Combining substance-specific data and database information. *Environ. Toxicol. Chem.* 22:1387-1393.

Sprague, J.B., 1969. Measurement of pollutant toxicity to fish I. Bioassay methods for acute toxicity. *Water Res.* 3:793-821.

Yours sincerely,

L.S. McCarty, Ph.D.

Michael Newman
Initial Responses to Questions

Charge Question 1.1* Please comment on the use of the *Guidelines for Ecological Risk Assessment* as an essential and relevant organizing framework for development of science-based criteria for the protection of aquatic life and aquatic-dependent wildlife. Does the SAB have any specific recommendations on how to improve or clarify the generic conceptual framework diagram?

Their use is consistent with trends in other aspects of EPA's functioning and actions being taken in the EU. I have no problem with this context.

Charge Question 1.2* Please comment on whether the proposed criteria types and the scientific focus for each criteria type are logical and scientifically valid for developing a holistic and integrated criteria framework.

Some of the proposed criteria types and the scientific focus for each are logical and scientifically valid. Those requiring refinement or modification are described below. Provided that the changes described below are made, the result should be scientifically valid for the stated purpose. In addition to the points made below, I suggest that mixtures be handled using classic models associated with mode of action and not using methods based on deviations from the (concentration) additivity. Current approaches derived from methods associated with additivity (such as the TU approach) are incorrect and should be de-emphasized.

Charge Question 2.1* Please comment on whether the kinetic toxicity models being considered by EPA are scientifically appropriate for use in deriving water-based criteria.

Although moving in an excellent direction, the proposed "kinetic toxicity model" approach has significant problems that require change before it is scientifically appropriate. This section requires inclusion of a broader vantage and incorporation of input from this SAB panel prior to being appropriate.

The major problems are technically resolvable because most seem to arise from unfamiliarity with the broad literature and a consequent lack of appreciation for the substantial work already done in other fields, especially the medical sciences. One publication (Mancini 1983) is cited repeatedly as the one that "pioneered" (page 12 of the Overview of Proposed Revisions of Guidelines for [Protection of Aquatic and Aquatic-dependent Wildlife]) this "new kinetic-based modeling approach" (page 24 of Final Report on Summary of Proposed Revisions). In reality, the approach is well-established in medical sciences, and Kooijmann and Bedaux (1996) in The Netherlands have a DEBTox program that already implements major components of the approach. (DEBTox is currently a controversial approach in the EU.) Three major issues require resolution prior to this approach being acceptable:

- (1) an intermediate option involving toxicodynamic (survival) analysis is required,
- (2) the document should be rewritten to incorporate the extensive, and often classic, literature covering this type of modeling, i.e., toxicokinetic-toxicodynamic modeling, and
- (3) outstanding issues to be resolved should be identified, discussed in the context of available approaches, and prioritized.

Each of these issues is discussed below.

- (1) An intermediate step involving solely toxicodynamics is required

The proposed approach will require acceptance by stakeholders, the regulated and regulated communities, groups within the general public, and the general scientific community. Success in getting any group to adopt a new idea depends on five qualities of the proposed innovation (Rogers 1995): (1) relative advantage over other approaches, (2) compatibility with existing “values, past experiences, and needs of the potential adopters,” (3) the degree to which a new approach is perceived as difficult to understand or apply, (4) the degree to which the new approach can be experimented with prior to full acceptance, and (5) the degree to which the results of the new approach are clearly observable to the adopters.

“Innovations that are perceived by individuals as having greater relative advantage, compatibility, trialability, observability, and less complexity will be adopted more rapidly than other innovations.” (Rogers 1995, page 16)

The proposed toxicokinetic-toxicodynamic approach scores poorly in my opinion relative to several of these characteristics. What one gets from the change is clear; however, the approach is very different from that presently applied. Toxicokinetic-toxicodynamic models take considerable effort to understand unless one treats them as simple black boxes. I assume that EPA does not want to treat these models as black box models. Yet, when examined closely, the actual toxicokinetics will not be considered in the proposed approach except in structuring the model to which the data will be fit. It will not be measured nor will associated assumptions be checked. There seems to be no intermediate or bridging approach although one will be proposed below. Lastly, the practical advantage will not be obvious relative to other approaches.

An intermediate “toxicodynamic” step is suggested for implementation in a tiered manner with the full toxicokinetic-toxicodynamic approach proposed here. Precedence has already been set for this suggestion of applying toxicodynamic models (i.e., survival or failure time models) and relating the results to conventional toxicity test and population models in the EPA ECOFRAM documents for aquatic biota.

In the proposed revised criteria documents, these toxicodynamic (survival or failure time) models were discussed very briefly and concluded to be statistical models of marginal utility. They are judged insufficient because they simply quantify mortality dynamics.

This is a very superficial conclusion. The population matrix models advocated later in the proposed criteria documents would be rejected according to these same rules because they only quantify birth, death and stage transition rates. They do not get into the bioenergetics and life history mechanisms of population change. Obviously, that would be inappropriate. Also, the proposed toxicokinetic-toxicodynamic model has a very significant “curve fitting” aspect to it. Indeed, the toxicokinetics will only be used to formulate the model that is then fit with mortality data. The toxicokinetic parameters will not be measured: they will be estimated from mortality data. The toxicokinetics model assumptions will not be checked during application. However, the simple model given in the “Proposed Revisions to Aquatic Life Guidelines – Water-based Criteria” is a one compartment, first order model with one uptake and one loss term.¹ This is certainly a model of computational convenience in which uptake clearance and elimination rate constants are simply summary statistics reflecting the overall contributions of the many complex physiological, organ, and cell processes resulting in uptake and elimination. The one compartment is a mathematical one, not a physical one. No metazoan is one physically uniform compartment within which the contaminant is instantly and uniformly distributed and displays no hysteresis. Completely mechanistic models that include all of these factors are called physiologically-based pharmacokinetic-pharmacodynamic (PBPK-PD) or physiologically-based toxicokinetic-toxicodynamic (PBTK-TD) models. PBPK-PD models are too difficult to parameterize for use here. So, the selection of the “kinetic” model alone is not adequate and is based in the false distinction of it being a mechanistic model and the survival models being merely statistical models. In reality, they both could be extremely useful if used insightfully and in balance with the answer being sought. Because toxicodynamic and the one compartment, 1st order toxicokinetic-toxicodynamic models are both partially mechanistic and partially statistical models, and an intermediate bridging model will likely be needed to get rapid implementation of the new approach, a tiered approach is proposed here. The underlying principle should be to develop a model of sufficient complexity to give useful predictions and then make it no more complex than necessary. Models would be of a form that, if needed, one could be extended into the form of the other easily. For some applications, a toxicodynamic (survival) model is sufficient and, for some others in which an understanding of toxicokinetics is needed and possible, a full toxicokinetic-toxicodynamic model should be applied. There is an enormous literature on relevant toxicodynamic (failure or survival) models (e.g., Cox and Oakes 1984, Marubini and Valsecchi 1995, Miller 1981,) that can easily stand alone or be expanded to include toxicokinetics.

Another reason for incorporating this intermediate survival model use is that accurate application of a toxicokinetic-toxicodynamic might not be possible when certain information is lacking or too costly to obtain. Examples include the following cases: risk of death is not directly related to the concentration in the body at that moment, a simple first order bioaccumulation model is demonstrably incorrect, induction of detoxification is significant, general stress or wasting determines time to death, previous exposure influences risk of death during a subsequent exposure, and uncertainty exists about the

¹ The uptake term was incorrectly defined in the proposed approach as a rate constant when, in fact, it is a clearance rate. See Barron et al. (1990) for details

relative contribution of individual qualities versus stochastic qualities to the risk of death at any moment. The discussion provided in support of the model does not include much of the published literature that addresses these issues, suggesting that these issues are not being seriously considered at this point. Specifics will be given below to address the question of whether these issues are important.

The assumption that there is a critical body burden that, if exceeded, results in death can be valid in some cases, but certainly not all cases. In some cases, metabolites are the compounds causing an adverse effect, not the parent compound. Some toxicants such as metals become sequestered in tissues and their levels are irrelevant to the level of effect occurring due to internal exposure. In some cases, the cumulative exposure (area under the concentration-time curve) might best reflect the risk of death, not the instantaneous body concentration. The amount in the entire body or single compartment proposed here might or might not be a good measure of the amount of toxicant available to have an effect at the site of action. Indeed, as stated in the quote below from a toxicology textbook, the concentration in a compartment such as the blood, is often the best concentration to relate to effect because the toxicant in the blood or plasma is constantly moving into the various organs where an effect may manifest.

“Once in the bloodstream the substance will distribute around the body and be diluted by the blood. Although only a small portion of the compound in the body may be in contact with the receptor or target site, it is the distribution of the bulk of the compound which governs the concentration and disposition of that critical proportion. The plasma concentration of the compound is therefore very important, because it often directly relates to the concentration at the site of action.” (Timbrell, 2000)

There will be cases in which the immediate damage to an organ such as the gill (oxidative damage of tissue by a metal), lung (nitrous oxide damage) or intestinal tract (denaturation of proteins after phenol ingestion) will be so quick that the internalized concentration is not a good indicator of toxicant concentration resulting in death.

The question of whether a one compartment, first order model with one uptake and one elimination component is a good one as a general model is easily answered. One simply needs to go to the literature to see how many times a more complex model was required to define the bioaccumulation of a toxicant in an aquatic or aquatics-related species. So many examples exist that I will not waste space here listing examples.

Many published studies, including those focused on biomarkers, demonstrate that induction of detoxification mechanisms occur often in the field. Such induction is not included in the proposed model but will influence transformation, deposition, and elimination kinetics. Equally important, when transformation results in activation of toxicant effect, the result will be toxicokinetic-toxicodynamic changes not accommodated by the proposed model.

Toxicants can contribute to the general wasting as demonstrated in Selye's general adaptation syndrome (GAS) which is a well-established theme in the literature. It is used to interpret many aspects of the effects of toxicants.

Previous exposure, as in the case of toxicant pulsed exposures, can influence the relationship between concentration and rate of mortality. In some cases, acclimation occurs and the risk decreases, e.g. fish exposure to dissolved copper (Chapman 1985). More recently, the adverse influence of previous exposure duration and concentration, and time between pulses have been quantified (Zhao and Newman, in press). The proposed model is inadequate for addressing these issues.

The question of whether the risk of dying during an exposure is influenced strongly or minimally by qualities of the individual is discussed in the proposed guidelines document. However, the discussion was not adequate and was written without consideration of classic and recent literature in the area. As this issue is relevant here, it will be discussed with more detail to illustrate how the proposed approach alone is insufficient to move ahead effectively.

The question in classic toxicology about how much individual qualities influence the form of toxicity models was addressed beginning in the 1930's by some extraordinary toxicologists. Yet no mention of this classic work is made in the documents supplied to me. While formulating the probit approach, Bliss (1935) proposed that individuals had distinct doses or concentrations above which they die and that the distribution of these "individual effective doses" (IEDs) is lognormal in most populations. The IED concept was established early in toxicology and reiterated in several early books on the subject. While advocating the logit approach, Berkson (1951) tested the IED concept and proposed a stochastic model for mortality curves instead. The two concepts, inaccurately named the "Deterministic Process Model" and "Stochastic Process Model"² in the document being reviewed are simply the IED and stochastic model concepts. These concepts and correct names are mentioned in a recent ecotoxicology textbook (Newman and Unger 2002), a book referenced in the report being reviewed (Newman 1995), and in the title of a recent research publication (Newman and McCloskey 2000). Ignoring issues already addressed in the existing literature and renaming concepts that already have names confuses the process and gives the false impression that EPA is unaware of the pertinent literature as it formulates important criteria. It delays linkage with existing, valuable work during incorporation of these concepts and methods into the criteria.

² On page 13 of the Proposed Revisions to Aquatic Life Guidelines – Water-based Criteria, this "Stochastic Process Model" is described as having a "fundamental variable", the hazard rate – "the probability of death per unit time for surviving individuals." Although I have made this error myself and sympathize with their error, they should change the text so that hazard rate is not defined as a probability. Because hazard rate can be larger than 1, it obviously is not a probability. An accurate definition can be obtained from any of the books cited below. The authors also state that the approach is parallel to that of modeling radioactive decay. It is only parallel if the hazard is constant and defined by an exponential function.

- (2) The proposed approach should be extensively rewritten to incorporate correctly the extensive, and even classic, literature that exists relative to this type of modeling, i.e., toxicokinetic-toxicodynamic modeling.

As discussed above, the description of the proposed approach does not incorporate classic and current important literature and basic concepts. The treatment of straightforward toxicodynamic (survival) models is inadequate and, as a result, neglects an extremely useful and potentially bridging approach between existing methods and new methods. For example, very good estimates of LC50 and other conventional toxicity metrics can be generated from survival models (e.g., Newman and Aplin 1992, Newman 1995). Survival modeling of some current toxicity test data sets can be incorporated directly into demographic models such as those proposed herein (Newman et al. 1994, Newman and McCloskey 2002). Numerous publications now exist that implement survival methods in ecotoxicology (see many references below).

The shortcomings of the current document reduces the defensibility of an otherwise good approach. It also will delay adoption of the new approach.

- (3) Outstanding issues to be resolved should be identified and prioritized.

In this document and the published literature numerous outstanding issues to be overcome are discussed. These include those associated with constant and variable (pulsed) exposure scenarios. They also include the potentially high mortality occurring after exposure ends (Newman and McCloskey 2000, Zhao and Newman 2004). Other issues include effects of mixtures (competing risks) and important covariates such as water quality and characteristics of individuals within the exposed populations. Perhaps a table listing the various issues and how the conventional, survival (toxicodynamic), and full toxicokinetic-toxicodynamic approaches resolve or fail to resolve each issue.

Charge Question 2.2.* Please comment on whether the population models being considered by EPA are scientifically appropriate for use in deriving water-based criteria.

There are two aspects to any answer to this question. First, is the issue of whether or not the population approach described here is appropriate. The second is whether or not the linkage of the toxicokinetic-toxicodynamic model to the population approach described here is appropriate.

- (1) Is the population approach appropriate?

The population approach described in the documents appears to be sound and forward looking. Some points will be made about specific issues in the spirit of enhancing the acceptance in use of the approach. The points will include details of implementing the approach as described and important shortcomings associated with population behavior that will not be addressed by the approach, and important shortcoming that will not be addressed relative to predicting population viability under toxicant stress within an ecological community.

Since populations are often metapopulations, is there consideration of expanding this approach to accommodate metapopulation-associated issues? This would seem an important issue to discuss.

Page 16 of the Proposed Aquatic Life Guidelines – Water-based Criteria document shows a stage-based (Lefkovitch) matrix approach to demographic projection.³ This seems the best choice but some species or situations might be better addressed with age-specific (Leslie) matrix approaches. I don't quite understand why both are not described and considered.

The approach does not consider stability criteria. Is this irrelevant? Isn't there a problem if the r remains above 0, the theoretical carrying capacity remains above some minimum, but the oscillations normally seen in some populations are influenced such that the risk of falling below the minimum through time is greatly increased? Are dynamics not to be considered in the density-independent approach? Some populations are inherently "boom-and-bust" in their behaviors. Simkiss et al. (1993) found that cadmium influences blowfly population oscillations.

In Section 4.1, point (b), please note that population density is also affected by transition probabilities between stages.

On page 18, there is mention of a species population's carrying capacity. In important cases, the concept of "a" carrying capacity is not an appropriate one.

There is an emphasis on recovery time or population growth rate. This is something that I would expect with a fisheries or range management activity. However, I fail to see the relevance of recovery time here. I suppose that one could use it in combination with estimates of how long a persistent contaminant is projected to remain in the habitat. Although r may seem the most universal metric of population state, it can be deceiving or misinterpreted. For example, populations show their highest productivity at 50% harvesting: r will be highest when half of the individuals are harvested. There seems a simple universal metric that is not emphasized, i.e., the probability of a local extinction under a specific exposure scenario. This statement of population risk can be produced by Monte Carlo simulation using the proposed matrix approach. It is easily estimated, directly relevant, and understandable by most stakeholders. It is directly related to other uses of risk in ecological risk assessment.

Given the proposed approach, it is confusing to me that sensitivity and elasticity analyses are not included in more detail. There is discussion of life cycle tests that are focused on the most sensitive stage of an individual's life cycle. This is not necessarily the most sensitive stage relative to population persistence. Newman (2002) refers to this common misconception as the weakest link incongruity. Kammenga et al. (1996) clearly demonstrated the flaw in this reasoning using soil nematodes exposed to cadmium. Because life cycle tests are discussed repeatedly, much of the discussion relates to asymptotic populations, and this error is common, it makes sense to discuss the means of

³ If the matrix is generic, the term in the top row, second column should be F_2 , not 0.

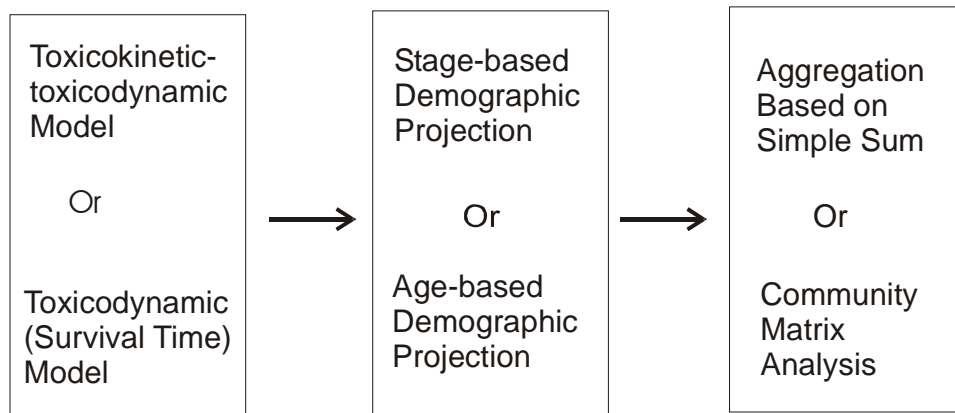
avoiding this error. In sensitivity and elasticity analysis, one uses the information discussed here to produce estimates of the sensitivity of r to changes in the various vital rates. Also the left eigenvector of the matrix provides very useful reproductive values estimates. Elasticity can then be estimated (e.g. pages 225-237 in Caswell (2001)). These are basic methods described in many books such as Caswell's much-cited textbook (2001) or Vandermeer and Goldberg (2003).

Finally, in interpreting results, it would be essential that changes in life history characteristics (life history strategies) and phenotypic plasticity be fully considered.

(2) Is the linkage of the toxicokinetic-toxicodynamic model to the population approach described here appropriate?

The linkage of the population model to the current toxicokinetic-toxicodynamic model is not appropriate but would be appropriate with the suggested changes. That linkage will be valuable in the next decade. The best approach is diagramed in the following figure in which sound ecological methods are melded as appropriate with toxicokinetic-toxicodynamic or toxicodynamic (survival) methods. Which approach is taken depends on the question or situation being addressed and the need for more or less detail. What internal metric of exposure to be used in the TK-TD model must be assessed and justified prior to application of this more involved modeling approach.

As appropriate for the question being asked ...



Charge Question 2.3.* Please comment on whether the proposal for aggregating effects across species being considered by EPA is scientifically appropriate for use in deriving water-based criteria.

On page 26, it is emphasized that the ultimate goal of the Aquatic Life Criteria is to protect the ecological community that is composed of species populations.

“It is important to note that this issue applies to the problem of protecting an assemblage as a whole, the role of the Aquatic Life General Use Criteria.”

On page 30, it is stated,

“That is, we could give each genus (rather than each species) equal weight, in the assemblage”

There are shortcomings that should be discussed more relative to viability of a population that is both impacted by a toxicant and members of its ecological community such as competitors, prey, and predators. The basic ecological premise that there are dominants and keystone species brings the second quote above into question. Before discussing these shortcomings, it should be stated that the approach advocated here is much better than the species sensitivity approach that neglects numerous fundamental concepts of ecology. The species sensitivity distribution approach should be removed from the Criteria.

In Section 5.2, the influence on the community of populations is treated as an issue of predicting what will happen to a population of independent populations. But a foundation concept of ecology is that communities are not independent populations that can be summed. The populations interact in very essential ways and classic methods are available for predicting the consequences of environmental changes on the viability of interacting populations. Some populations are more important than others because of their dominance or function. Newman (1995) discusses well-established competition or community matrices that emerge as extensions/expansions of the population models being advocated herein. As discussed by Newman (1995), there are very classical ecological means of assessing stability regions, i.e., species risk of being unable to co-exist with others in the community. Although the general adoption of such community matrix methods are not recommended here, they should be understood, used to provide insight during application, and applied in interpreting the results of the proposed methods.

Charge Question 2.4.* Please comment on whether the framework being considered by EPA for deriving water-based criteria is scientifically appropriate for use in deriving the criteria.

Yes. Provided that the changes described above are made, the framework will advance the Agency’s ability to protect aquatic and aquatic-related life.

References

Barron, M.G., G.R. Stehly, and W.L. Hayton, 1990. Pharmacokinetic modeling in aquatic animals, I: models and concepts, *Aquatic Toxicol.* 18, 61-86.

- Berkson, J., 1951, Why I prefer logits to probits, *Biometrics* 7: 327-339.
- Bliss, C.I., 1935, The calculation of the dosage-mortality curve, *Ann. Appl. Biol.* 22, 134-307.
- Caswell, H., 2001, *Matrix Population Models. Construction, Analysis, and Interpretation*, 2nd Ed., Sinauer Assoc. inc, Sunderland, MA, p. 722.
- Chapman, G.A., 1985, Acclimation as a factor influencing metal criteria, in, *Aquatic Toxicology and Hazard Assessment*, 8th Ed., ASTM STP 891, R.C. Bahner and D.J. Hansen, Eds., ASTM, Philadelphia, PA, pp. 119-136.
- Cox, D.R., and D. Oakes, 1984, *Analysis of Survival Data*, Chapman and Hall, New York, p. 201.
- Dixon, P. M. and M. C. Newman. 1991. Analyzing toxicity data using statistical models of time-to-death: An introduction. In: Newman, M.C. and A.W. McIntosh (Eds.), *Metal Ecotoxicology: Concepts and Applications*. Lewis Publishers, Inc., Chelsea, MI., pp. 207-242.
- Marubini, E. and Valsecchi, M.G., 1995, *Analysing Survival Data from Clinical Trials and Observational Studies*, John Wiley and Sons, New York, p. 413.
- Miller, Jr., R. G., *Survival Analysis*, John Wiley and Sons, New York, p. 238.
- Newman, M.C. and M. Aplin. 1992. Enhancing toxicity data interpretation and prediction of ecological risk with survival time modeling: An illustration using sodium chloride toxicity to mosquitofish (*Gambusia holbrooki*). *Aquatic Toxicol.* 23:85-96.
- Newman, M.C., M.M. Keklak and S.M. Doggett. 1994. Quantifying animal size effects on toxicity: A general approach. *Aquatic Toxicol.* 28:1-13.
- Kammenga, J.E., M. Busshers, N.M. Van Straalen, P.C. Jepson, and J. Baker, 1996, Stress induced fitness is not determined by the most sensitive life-cycle trait, *Functional Ecol.* 10: 106-111.
- Kooijmann, S.A.L.M. and J.J.M. Bedaux, 1996, *The Analysis of Aquatic Toxicity Data*, VU University Press, Amsterdam, The Netherlands, p. 149.
- Newman, M.C. and J.T. McCloskey. 1996. Time-to-event analysis of ecotoxicity data. *Ecotoxicology* 5: 187-196.

- Newman, M.C. and P.H. Dixon. 1996. Ecologically meaningful estimates of lethal effect on individuals. In: Newman, M.C. and C.H. Jagoe (Eds.), *Ecotoxicology: A Hierarchical Treatment*. CRC/Lewis Publishers, Inc., Boca Raton, FL, pp. 225-253.
- Newman, M.C. and J.T. McCloskey. 2000. The individual tolerance concept is not the sole explanation for the probit dose-effect model. *Environ. Toxicol. Chem.* 19:520-526.
- Newman, M.C. and J.T. McCloskey. 2002. Applying time-to-event methods to assess pollutant effects to populations. In: Crane, M., M.C. Newman, P. Chapman, and J. Fenlon (Eds.) *Risk Assessment with Time-to-event Models*. CRC Press LLC.
- Rogers, E.M., *Diffusion of Innovations*, 4th Ed., The Free Press, Collier Macmillan Publishers, New York, p. 519.
- Simkiss, K., S. Daniels, and R.H. Smith. 1993, Effects of population density and cadmium on growth and survival of blowflies, *Environ. Pollut.* 81, 41-45.,
- Timbrell, J., *Principles of Biochemical Toxicology*, 3rd Ed., Taylor & Francis, Philadelphia, PA, 2000, p. 394.
- Vandermeer, J.H., and Goldberg, D.E., 2003. *Population Ecology, First principles*, Princeton University Press, Princeton, NJ, p. 280.
- Zhao, Y. and M.C. Newman. 2004. Shortcomings of the laboratory derived LC50 for predicting mortality in field populations: exposure duration and latent mortality. *Environ. Toxicol. Chem.* 23: 2147-2153
- Zhao, Y. and M.C. Newman. Effects of exposure duration and recovery time during pulsed exposures. *Environ. Toxicol. Chem.* (in press)

Date: 19 September 2005

To: Dr. Thomas Armitage

From: Robin J. Reash

Subject: SAB Aquatic Life Criteria Guidelines Panel
Initial Written Responses to Charge Questions

Here are my initial written responses to the charge questions that I was asked to comment on pursuant to the 26 August 2005 memo from you. Please note that my comments may be supplemented based on comments from other panel members (and other participants attending the 21 September meeting).

Charge Question 1.1 – Please comment on the use of the Guidelines for Ecological Risk Assessment as an essential and relevant organizing framework for development of science-based criteria for the protection of aquatic life and aquatic-dependent wildlife. Does the SAB have any specific recommendations on how to improve or clarify the generic conceptual framework diagram?

EPA's ecological risk assessment guidelines ("ERA Guidelines") define a process whereby management decisions can be evaluated based on the robustness of underlying scientific information and the level of inferred risk to ecological receptors. As a general framework for directing the process of how water quality criteria are developed, I believe the ERA guidelines are an appropriate mechanism. There are advantages of using the ERA guidelines: 1) EPA has spent considerable time developing the essential inputs of what constitutes an effective decision-making process; 2) the ERA paradigm, while probably underutilized by stakeholders on a national scale, has been used in several site-specific risk evaluations for a number of different pollutants and exposure scenarios; 3) theoretically, the ERA process is iterative (not deterministic), and is designed to allow shifts in environmental protection (management) decisions commensurate with new information; and 4) the ERA process explicitly considers uncertainty for all input variables. In short, the ERA process is designed to provide decision-making flexibility based on the level of protection desired and the underlying certainty of scientific information.

There are some differences between the ERA paradigm and how aquatic life criteria are developed (at least historically) that EPA should be aware of. First, the ERA guidelines support management decisions that are highly site-specific. There are concise boundaries to a typical ERA evaluation: geographic area of exposure, ecological receptors,

contaminant occurrence and bioavailability, etc. The linkage between management decisions and risks are more certain, i.e., the universe of potential adverse risks is limited.

The development of aquatic life criteria, in contrast, is largely inferential. Historically, EPA has developed aquatic life criteria for generic (universal) application, while providing sound technical procedures for site-specific modifications. In the Risk Characterization stage for criteria development, the selection of endpoints (for protection of waterbody use protection) can be subjective. In a traditional ERA process, this decision is considerably more intuitive because there is more certainty about what adverse effects would occur based on site-specific ecosystem attributes. Historically, EPA has based aquatic life criteria on non-controversial endpoints: protection against mortality, decreased growth, and impaired reproductive potential. Selection of endpoints for national criteria (some of which could be subtle sublethal ones) will require considerable judgement. For development of new aquatic life criteria, EPA faces challenges in the Risk Characterization phase.

I strongly support using the Problem Formulation phase as the first step of developing aquatic life criteria. ***The key outcome is prioritization.*** Traditionally, EPA has not fully developed the elements of Problem Formulation in developing nationally-recommended aquatic life criteria. There is a need for EPA to enhance aquatic life criteria documents with more discussion of contaminant biogeochemistry, sources of a pollutant (natural and anthropogenic), known examples of ecological impacts, and interactions with other pollutants. In the 2004 Draft Aquatic Life Criteria for Selenium document, EPA made improvements in addressing these issues. For example, there was good discussion on factors causing site-specific variations in contaminant effect/no effect. In the Problem Formulation stage, EPA needs to make a good technical case that a specific aquatic life criterion (whether new or revised) is justified. In many cases new criteria are justified because new, significant information is available. Presumably, the limiting factor affecting HOW MANY criteria can be reviewed – at any given time – is funding. The Problem Formulation phase could provide key information supporting management decisions of funding.

The final phase of the ERA process (Risk Management) consists of weighing risks and benefits with all other factors (legal, social, economic, etc). In site-specific settings, the costs and benefits of management decisions have, in general, good resolution. What may be difficult for EPA is to “fit” the Risk Management phase – typically the final element of ERA evaluations – for development of nationally-recommended aquatic life criteria. I propose that EPA consider factors at the Risk Management phase that, typically, have not been evaluated in the development of aquatic life criteria. Some examples of these are listed below:

- Analytical feasibility and challenges: is it appropriate to issue aquatic life criteria that are below the Method Detection Level for the most sensitive EPA-approved analytical methods?

- Comparison to background levels: is it appropriate for EPA to issue aquatic life criteria that are less than typical background environmental levels?
- Pollutant speciation: if species-specific aquatic life criteria can be justified based on differential toxicity of forms (e.g., arsenic and selenium), what are the barriers that prevent implementation of these criteria at the state level?

In short, the challenge that EPA faces during the Risk Management phase of developing aquatic life criteria is addressing the issues stated on page 8 of the ERA Guidelines:

The wide use and important advantages of ecological risk assessments do not mean they are the sole determinants of management decisions; risk managers consider many factors. Legal mandates and political, social, and economic considerations may lead risk managers to make decisions that are more or less protective. Reducing risk to the lowest level may be too expensive or not technically feasible. Thus, although ecological risk assessments provide critical information to risk managers, they are only part of the environmental decision-making process.

If EPA truly intends to complete criteria evaluations using the ERA-based Risk Management phase, considerable discussion on the overall process (and factors to be considered) will be needed.

In summary, I support the framework of the ERA Guidelines to guide the overall process of developing aquatic life criteria. Scientifically defensible, protective aquatic life criteria can be the outcome of the ERA paradigm. EPA, however, will need to more fully discuss the decision-making process during the Risk Characterization and Risk Management phases during criteria development. Many of the considerations in the Risk Management phase of the ERA Guidelines are not, currently, assessed by EPA during the development of aquatic life criteria.

Charge question 1.2 – Please comment on whether the proposed criteria types and scientific focus for each criteria are logical and scientifically valid for developing a holistic and integrated criteria framework.

In one sense, EPA has made good progress on developing criteria that are “outside the box” of criteria that are typical outcomes of the 1985 guidelines. EPA’s fish tissue-based methylmercury human health criterion and proposed aquatic life selenium chronic criterion (also fish tissue-based) are good examples of the agency focusing on the most critical, appropriate media to regulate pollutants. There are several implementation issues that have not been resolved with these criteria, but the scientific justification of expressing these criteria as tissue levels (as opposed to water column concentrations) was strong.

I like the idea of allowing scientific knowledge direct how a criterion is developed and expressed. The three types of criteria being considered by the agency are appropriate, but additional types of criteria may be justified in the future. The goal of valid aquatic life criteria is to provide a sound technical foundation for preventing unacceptable adverse impacts, while not overprotecting (deliberately protecting against speculative or ecologically irrelevant endpoints or outcomes). Optimally, a given pollutant would have as many criteria types that are scientifically justified. EPA does state that a given pollutant could have both water-based and tissue-based criteria. I like this concept because *it provides flexibility on the application of criteria*. A major challenge facing EPA is developing aquatic life criteria that are, at least, somewhat representative of all waterbody types. EPA is aware, however, that each waterbody has unique exposure settings. The development of paired criteria, expressed in different media, is a first step in allowing a more appropriate fitting of criteria to exposure settings. There are some exposure settings where tissue-based criteria (or the traditional chronic criteria) are probably not appropriate because exposure assumptions of these criteria are clearly not present. In contrast, tissue-based criteria are clearly needed for high exposure/high bioaccumulation potential settings, especially when regulating pollutant loadings using water-based criteria only may be underprotective.

I have more specific comments on water-based criteria in the following section, so a few remarks on tissue-based and taxon-specific criteria appear below:

The process of developing tissue-based criteria needs to provide for alternate assumptions (input variables) that can be used for state, region, or site-specific applications. For the purpose of developing 304(a) water quality criteria, it is proper for EPA to develop these criteria using default assumptions (often these are worst-case) and considering all valid toxicity data for all species. A tissue-based criterion, however, is an *estimate* of an accumulated pollutant concentration that should be protective considering exposure scenarios consistent with the underlying effects database. The criteria, however, need to identify all relevant input variables that can be utilized for state, regional, or site-specific applications. *EPA should lay the scientific foundation for tissue-based criteria that specifically encourage the tailoring of such criteria to settings having higher certainty of input variables*. Some of these considerations are:

- The presence or absence of certain taxa relative to the most sensitive species in the toxicity database;
- The structure of the food web or feeding ecology of the most sensitive species
- Chemical and physical characteristics of a waterbody (e.g., do these characteristics promote a high bioaccumulation potential?)
- The speciation and media flux of a bioaccumulative pollutant (e.g., is there a reasonable correlation between water column concentration and fish tissue concentration, or is an alternate biotic or abiotic compartment more appropriate?)

A final comment on water quality criteria development and implementation. My opinion is that EPA should strongly consider expressing 304(a) criteria in menu-type fashion. The goal here is to give states and tribes more choices, with choices being segregated by sound scientific rationale. Because I am most familiar with metals aquatic life criteria, I offer some examples of how EPA could provide more menu-type choices for these kind of criteria. These examples would not replace, or make invalid, the three kinds of criteria EPA is currently evaluating (water-based, tissue-based, taxon-specific based).

- Provide oxidation state-specific criteria. EPA made good progress in this area by proposing oxidation state-specific continuous maximum concentration (CMC) criteria for selenium. Other toxicants that should be evaluated for individual species criteria are arsenic, aluminum, and thallium. Clearly, the implementation of oxidation state-specific criteria is problematic from analytical, monitoring, and enforcement standpoints. EPA should, nonetheless, strive to express criteria as individual species if a sound toxicological rationale can be demonstrated. EPA could also issue criteria based on the “total metal” form, alongside the oxidation state-specific criteria. States and tribes can then choose which criteria to adopt based on a number of considerations.
- Provide multiple criteria based on the inclusion or exclusion of certain taxa. In some cases, there is good evidence that certain taxonomic groups are more sensitive than other groups within the toxicity effects database. Recent findings have shown that freshwater unionid mussels are particularly sensitive to ammonia and some trace metals. Mussel populations and communities are typically patchy in their distribution, their presence often limited by a few critical habitat features. EPA should consider providing dual or multiple criteria that are based on the presence or absence of certain taxa, where such taxa are demonstrated (as a group) to be particularly sensitive. Once again, states and tribes would have the choice to adopt one or both of the dual criteria.
- Provide criteria that are segregated by bioavailability-type adjustments. EPA should be commended by allowing states and tribes the flexibility to adopt metals criteria based on the dissolved portion of the metal. Specific conversion factors have been provided in previous water quality standards rulemakings. While some states have not yet adopted dissolved metals criteria, EPA should consider either issuing future metals criteria as dissolved only, or issue both dissolved and “total” criteria.

The development of the Biotic Ligand Model (BLM) presents both advantages and disadvantages from a criteria implementation standpoint. The BLM represents a much more mechanistic, scientifically valid framework for elucidating (and regulating) the true bioavailable metal, compared to a dissolved metals approach. Implementation of the BLM, however, is operationally problematic due the requirement of valid water quality input variables for the

model. Moreover, many stakeholders have cited concerns that the BLM approach, while sound scientifically, is limited to site-specific or waterbody-specific application only. While the validity of this concern may be argued, EPA should encourage the BLM approach as a valid option to regulate trace metals. As such, EPA should consider issuing both bioavailability-based criteria and traditional criteria (hardness-based, dissolved, or total metal). Once again, the goal of these recommendations is to encourage EPA to maximize the choices states and tribes can evaluate.

Charge Questions for Water-Based Criteria

Charge Question 2.1 – Please comment on whether the kinetic toxicity models being considered by EPA are scientifically appropriate for use in deriving water-based criteria

The use of a kinetic toxicity model approach would improve the current aquatic life criteria calculation assumptions of: 1) continuous exposure at “steady state” internal levels and 2) a generic toxicant “speed of action” for all contaminants. EPA’s central justification of a kinetic modeling approach (internal toxicant concentration being a function of competing rates of accumulation and depuration) is sound from a toxicological basis. EPA’s recommended criterion averaging time parameters have been criticized for lack of biological realism, but I believe the more important criticism is the generic assumption of “speed of action”. EPA’s recommended averaging time of 24 hrs for acute aquatic life criteria is based on measured response of test organisms to a fast acting pollutant (e.g., ammonia). A pollutant-specific “speed of action” estimate is clearly needed, and a Mancini-Breck type of kinetic model would be appropriate. ***A mechanism is clearly needed to define organism response to toxicants with differing modes and speeds of action. In addition, the associated averaging period and return frequency for criteria exceedance needs to reflect pollutant-specific modes of action and speeds of action.***

EPA’s desire to model the response of aquatic organisms based on a time-varying concentration pattern is appropriate in the goal of understanding organism response to realistic environmental exposures. Time-varying exposure scenarios likely predominate in most wastewater discharge/receiving stream settings. While the concentration or loading of toxicants from a point source can vary along many temporal scales, variation in receiving stream flow is the likely cause of temporally variable instream toxicant concentrations in most situations. It should be noted, however, that there are settings where a continuous, or quasi-continuous, exposure regime is the norm. Effluent-dominated streams and point source discharges to lentic waterbodies are two examples. In these settings, a continuous “steady state” exposure assumption may be appropriate.

I believe that a deterministic process model would be more technically valid compared to a stochastic model approach. Variability in sensitivity of individuals within a population is one of the cornerstone principals of aquatic toxicology. Moreover, the empirical

evidence for deterministic population response seems to be much greater than a stochastic response.

EPA is proposing to replace the existing paired acute/chronic criteria with a single criterion value (having a specified averaging period and exceedance return frequency). A single criterion value would have implementation advantages concerning the establishment of effluent limits and TMDLs, and in permit compliance. My concern is that some exposure settings may not be relevant to the long-term assumption of population modeling. Specifically, the proposed population model is meant to estimate population size over several generations. The exposure of aquatic biota to some pollutants, in contrast, may be limited to a fraction of the organism's lifespan. Certainly, organism response to relatively short exposure durations can be the outcome of kinetic modeling. EPA should nonetheless consider how site-specific adjustments could be made, of resulting criteria, based on markedly different site-specific exposure regimes (relative to the long-term regime postulated by the proposed population modeling).

I believe there are some situations where dual acute/chronic criteria should be applied. In some states, the application of both acute and chronic criteria do not apply to some use designations. For these use designations, protection against acute effects only is provided due to irretrievable man-induced conditions or the presence of other factors that prevent the attainment of "fishable, swimmable" uses. EPA may want to consider the value of having dual acute and chronic criteria as an option for states. At minimum, states and tribes should be consulted regarding possible implementation difficulties with using a single aquatic life criterion to protect against short and long-term exposure effects.

EPA should consider modeling highly episodic exposure settings (short duration of high exposure separated by long periods of little or no exposure). Storm-induced exposures, for example, may or may not elicit adverse effects upon receiving stream biota. Many states, however, regulate stormwater discharges using traditional "continuous exposure" aquatic life criteria. Kinetic modeling results (and resulting criteria) for episodic exposure scenarios would be more appropriate for regulating storm-induced discharges.

Recent toxicological findings have identified variable toxicant sensitivities between laboratory populations that have had no prior exposure to a toxicant ("naive populations") and those populations that are exposed to low (i.e., less than chronic threshold) levels of the same toxicant. The issue of pre-acclimated tolerance, and the significance to development of ambient aquatic life criteria, should be reviewed by EPA. At minimum, EPA should clarify that kinetic modeling outputs are, or are not, assumed to represent "naive exposure" populations.

Charge Question 2.2 – Please comment on whether the population models being considered by EPA are scientifically appropriate for use in deriving water-based criteria.

EPA is considering the use of a life stage-structured, matrix, density-dependent population model to supplement the kinetic modeling outputs in the development of aquatic life criteria. The kinetic modeling outputs would define the pattern of organism response to a stressor using an intermittent exposure scenario. The population model would embed the organism response patterns (for specific life stages) to predict population-level responses (survival, fecundity, reproductive success). The model would explicitly include parameters for the rate of recovery (based on kinetic modeling results) and compensation (a response to certain individuals being “culled out” by toxicant stress).

The use of population models to elucidate “acceptable” pollutant exposure regimes has some technical advantages. The traditional criteria derivation procedure is deterministic regarding how results for key endpoints are assessed: individuals from a population either survive or do not survive, and lowered reproductive potential and/or growth is a non-varying population “constant” whereby such affects do not occur at pollutant concentrations just below this level. Most population models have some basis in empirical data (i.e., observations of field populations). For example, density-dependent compensation has been observed in many fish populations subjected to significant fishing pressure, or when these populations are allowed to recover as a result of harvest restrictions. While the *occurrence* of compensation is generally supported in the technical literature, considerable debate is focused on the magnitude of, and biological processes underlying, the compensation. In the sense that population modeling more closely approximates real-world population responses to chemical stressors, I believe this approach has technical advantages to the existing criteria derivation assumptions.

I have some concerns and cautions regarding the use of population models. Population modeling can be an effective method to predict comparative responses of different populations to a stressor. The accuracy of population modeling outputs is proportional to the accuracy of input parameters. ***Thus, scientific confidence in parameterization is critical.*** The effectiveness of population modeling is reduced by using speculative population parameters (e.g., parameters that are simply assumed, based on unrelated taxa, or have considerable variation in measured values). EPA, therefore, needs to evaluate the rigor of information (data quality and quantity) that would constitute acceptable modeling parameters.

Many populations of aquatic biota display significant fluctuations in population size or density, often caused by interannual variability in abiotic variables. In the field, large-scale variability in population size can explain the inability to detect (and identify) adverse effects caused by anthropogenic activities. EPA should consider embedding model parameters that account for significant temporal variability of population size. Side-by-side modeling scenarios (with and without these adjustments) would provide useful information. It may be argued that natural variability in population size is an element, or kind, of compensation. I disagree with this position, since compensation is a result of density-dependent processes. Interannual variability of population size can be a factor in a population's rate of recovery following toxicant-induced stress.

I agree that EPA should attempt to incorporate life history stage-specific toxicity data when such information is available. Caution should be used, however, in extrapolating sensitivities of individual life stages across species or taxa. In addition, it should not always be assumed that the early (juvenile) life stage is more sensitive to the adult stage, for a given toxicant. Exceptions to this generalization do occur.

Charge Question 2.3 – Please comment on whether the proposal for aggregating effects across species being considered by EPA is scientifically appropriate for use in deriving water-based criteria

Protection of an entire assemblage (community) is certainly desirable, but there are many processes that the TEA outputs would be unable to incorporate. Biological interactions (e.g., predation, trophic competition, species displacements) can be important regulating mechanisms of assemblages. I do recognize that EPA is not attempting to model, or incorporate, such processes in the assemblage modeling. In terms of an aggregation model attempting to define the level of impairment (such that a level less than this would not cause impairment) for a single, specific stressor (toxicant exposure), I think the general approach is 1) conceptually valid, and 2) would be consistent with some goals of the existing aquatic life criterion derivation procedure.

I have two principal concerns with usage of a model that seeks to define the level of impairment for an assemblage of species. First, EPA needs to consider how results for site-specific biological assessments (community level) are reconciled with monitoring information suggesting that exceedances of assemblage-based criteria are impairing an aquatic life use designation. While EPA may believe that such considerations are beyond the scope of this new aquatic life criteria framework, real-world experience does indicate that biological communities are often not impaired even when chemical-specific criteria indicate impairment.

My second concern regards how the structure of the toxicity database assemblage is considered when criteria for assemblage protection are developed. The proposed TEA model does not seek to account for the redundancy of assemblage ecological function within the toxicity database. In the proposed approach, an adverse impact upon one functional level (primary consumers, for example) would be “equivalent” regardless of whether one, or multiple, primary consumer taxa are within the toxicity database. If the toxicity effects database contains toxicological endpoints for five or six taxa within a functional group, the long term “loss” of one, or two, taxa (due to toxicant stress) would not impair the overall function of the community. In contrast, the “loss” of a sole taxa for one functional group would be expected to cause significant impairment to assemblage function. ***In short, EPA needs to assess the sensitivity of an assemblage based on the representation of different ecological functional groups within the toxicity effects database.***

Charge Question 2.4 – Please comment on whether the framework being considered by EPA for deriving water-based criteria is scientifically appropriate of use in deriving the criteria

Clearly, the Water-Based Criteria Subcommittee has made extensive progress in evaluating how to improve the scientific validity of aquatic life criteria for pollutants that typically elicit adverse effects through water column exposures. Overall, the proposed new framework represents a significant improvement in the scientific foundation of developing protective (but not overprotective) aquatic life criteria, compared to EPA's existing framework (1985 guidelines). My major comments, and concerns, regarding the proposed new framework are identified in each charge question.

I suggest that EPA proceed in a stepwise, cautious manner concerning a possible wholesale overhaul of the existing aquatic life criteria derivation paradigm. Kinetic and population modeling is technically appropriate, but good empirical results will need to be obtained. I suggest that EPA select a few pollutants (having differing modes of action), conduct robust kinetic and population modeling for a core group of test organism species, evaluate the results, and contrast the criteria outputs with criteria derived using the traditional procedures. EPA may then want to ask the question, "which of these criteria makes more biological sense"?

Comments on the Scope of the Proposed Framework for Revising the Aquatic Life Water Quality Criteria

Daniel Schlenk
UC Riverside

1.1 Please comment on the use of the *Guidelines for Ecological Risk Assessment* as an essential and relevant organizing framework for development of science-based criteria for the protection of aquatic life and aquatic-dependent wildlife. Does the SAB have any specific recommendations on how to improve or clarify the generic conceptual framework diagram?

Utilizing the *Guidelines for Ecological Risk Assessment* is an excellent first step in organizing a framework for the development of criteria for the protection of wildlife. It is critical to blend aspects of exposure and adverse effects in setting realistic criteria for protecting populations of aquatic organisms. Discussion of methods to better evaluate acute and population level changes is an important improvement over the 1985 document. While the development of these criteria are a logical progression toward better evaluation of single stressors, it is recommended that the impacts of stressor mixtures be considered for the development of future criteria for aquatic organisms. The topic of mixtures of chemical and non-chemical stressors was largely avoided in most of the proposed revisions.

A great example would be the relative protecting effects of Selenium on Mercury toxicity potentially resulting from consumption of contaminated aquatic organisms by wildlife (or humans). If a tissue-based standard is set for mercury alone without considering the protective effects of Selenium, the criteria will likely be overprotective.

It would seem that the effects of any given toxicant on a population of aquatic organisms are going to be influenced by numerous other stressors which may or may not be of a chemical nature (i.e. habitat destruction, hypoxia, sediment loading) necessitating site specific criteria. Understanding the modes of action of each stressor and the impacts of these upon a specific toxicant will be crucial in setting “scientifically defensible numeric criteria”. Such a procedure is being formulated with regard to pharmaceutical and personal care chemical mixtures, using mode of action analyses to determine additive, synergistic and/or antagonistic interactions between potentially hundreds of compounds. Thus, it is recommended that more of a site-specific approach be outlined in the conceptual framework within the problem formulation step.

1.2 Please comment on whether the proposed criteria types and the scientific focus for each criteria type are logical and scientifically valid for developing a holistic and integrated criteria framework.

The three criteria targeted in the proposed criteria and the scientific focus (utilizing the ERA approach) are logical and valid for developing a holistic and integrated framework. Implementing a toxicokinetic evaluation method coupled to population assessment will likely provide a more efficient and ecologically relevant endpoint toward criteria development. Development of criteria for diet-derived bioaccumulative stressors is valid particularly for compounds that magnify through trophic levels. And, finally, consideration of criteria for species of special concern (i.e. endangered) is necessary as these organisms are often receptors of interest at various locations impacted by stressors. As mentioned above, however, it is recommended that a higher degree of attention be directed toward impacts of chemical and non-chemical stressor mixtures and that the framework be focused more upon site specific analyses

Taxon-Specific Criteria

Please comment on the considerations for problem formulation outlined in the proposed framework for deriving taxon-specific criteria, specifically whether it will lead to scientifically-defensible numeric criteria.

Developing taxon-specific guidelines are certainly necessary for endangered species or species of special concern. Gearing the problem formulation process toward protection of said species is a sound “first step” in providing potential protection. While the topics provided are indeed necessary to provide sound assessment endpoints for problem formulation. The issues are not clearly defined in the document. Therefore, based upon the text provided, it is unclear whether “scientifically-defensible” numeric criteria can be obtained. For example, a better description of how “effects to habitat” will be used with regard to taxon-specific criteria is warranted.

Of the approaches outline for addressing surrogacy and gap analyses with regard to special status taxa, are there improvements to these tools that would provide more scientifically defensible numeric criteria where specific data are not available? Are these tools adequate for developing defensible numeric criteria? What other tools are available to provide defensible criteria when there is an absence of toxicological data for a specific pollutant and taxon?

Most of the suggested approaches appear to dwell upon acute endpoints utilizing either ICE or SSD analyses. Given the appropriate data, it is possible “scientifically defensible numeric criteria” might be attained for acute endpoints. It may be possible that SSD analyses may be potentially capable of utilizing probabilistic evaluations of chronic endpoints which are likely more crucial to population level effects given the appropriate degree of data.

A specific guideline or protocol for the reduction of data gaps appears to be lacking from the document. It would seem more helpful if a decision tree analysis be implemented allowing regulators to pursue alternative strategies. For example, if data is lacking for a particular species, then pursue data on another species in the same genus. If this is

unavailable , then utilize data from a species that has similar life history (i.e. niche) and reproductive pathways (i.e. populations of animals with slower reproductive pathways tend to be more sensitive to impairment). However, an these criteria would possess a much higher degree of uncertainty and impair the “defensiveness of the criteria”. It is certainly not wise to utilize a different organism from a different family or having a different reproductive strategy to set criteria for a species of special concern. Generally, organisms of similar genetics would be expected to respond in similar ways. However, it should be noted that organisms of the same species located at historically polluted sites respond significantly different to stressors than populations of the same species from lesser impacted location. Hopefully advances in genomic analyses of each strain may help in targeting identifying “toxicant responses” in various taxon at specific sites which will reduce uncertainty in risk assessments and eventually provide better evaluations of population stability.

Proposed Revisions to the Aquatic Life Criteria Guidelines

Initial Responses to Charge Questions WA Stubblefield

Charge Question 1.1 Please comment on the use of the Guidelines for Ecological Risk Assessment as an essential and relevant organizing framework for development of science-based criteria for the protection of aquatic life and aquatic-dependent wildlife. Does the SAB have any specific recommendations on how to improve or clarify the generic conceptual framework diagram?

The use of the Guidelines for Ecological Risk Assessment (ERA) as a framework for the development of science-based water quality criteria is wholly-appropriate and consistent with both the state-of-the-science and current USEPA regulatory approaches. The risk assessment framework provides a structured method for evaluating the fate and effects of environmental contaminants and provides a consistent method by which to evaluate possible environmental risks. Figure 2 in the Consultation document describes the overall conceptual framework and adequately shows the theoretical relationships between the various key aspects of the model. However, the model is very generic as presented and could benefit from greater specificity and examples (in the text). Additionally, concern about sediment criteria values and their relationship to water-based criteria is somewhat unclear.

The ERA framework approach is an appropriate methodology for developing and implementing Criteria in a regulatory method. The Agency, however, may want to give additional consideration to the end-use of the criteria values and modify, as appropriate, the proposed ERA approach. For example, the classical use of AWQC values is as numeric guidance values that are, in most cases, adopted by States as State Standards that are in turn used to regulate the discharge (primarily point-source) of contaminants under the National Pollutant Discharge Elimination System (NPDES). The assumptions regarding the magnitude, duration, and frequency of criteria exceedances are to some extent predicated on assumptions developed for the NPDES program. However, AWQC values are frequently adopted (correctly or incorrectly) for other uses (e.g., site risk assessments, Natural Resource Damage Assessment [NRDA]) without consideration for the assumptions inherent in the criteria development process. The fact that criteria are not intended to be “bright line” values that if met will ensure no environmental consequences and if exceeded portend adverse environmental effects is often forgotten. For specific issues, such as control of stormwater contaminants, it may be appropriate to adjust both the fate (e.g., short-duration, high concentration pulse exposures) and effects (e.g., pulse toxicity test results, i.e., <48-96h exposures) assumptions—a flexible approach based on the ERA conceptual model will achieve this.

Charge Question 1.2. Please comment on whether the proposed criteria types and the scientific focus for each criteria type are logical and scientifically valid for developing a holistic and integrated criteria framework.

The three criteria types (water-based, tissue-based, and taxon-specific) are scientifically appropriate for the proposed framework and present an approach that is state-of-the-science, if not somewhat cutting-edge. A couple of brief points:

- It is unclear what role sediment criteria will play or what their relationship will be to water-based criteria. This should be considered.
- I am somewhat confused why it is that the taxon-specific criteria approach is separated from the water-based criteria approach. It appears that it could be adequately considered within the water-based approach.
- The Consultation document indicates a management objective that should address assessment endpoints that include:
 - "...quality and health of aquatic organisms",
 - "Size, sustainability, and resiliency of aquatic populations."
 - Sustainability, resiliency, diversity, structure, function, productivity, stability, and composition of aquatic communities and ecosystems." In addition to the normal survival growth and reproduction. I am unclear as to how the Agency would propose to quantitatively identify let alone incorporate these endpoints into a criteria model.

The USEPA should be commended for their efforts to incorporate current scientific thinking into the criteria development process. However, I think that we would be remiss if we did not acknowledge the fact that the 1985 approach (and its predecessor) has served us well and has gone a long way to improve environmental protection. The original authors of the extant criteria approach were clearly "forward-thinking" individuals who showed a great deal of perceptivity and creativity in developing the 1985 approach. Given the magnitude of the effort and costs that will be required to integrate the "new" criteria approach into the regulatory main stream, perhaps EPA would be wise to develop a "tiered" or "step-wise" research approach for development and implementation of the "new" criteria derivation approach while continuing to improve and modify the 1985 document. Clearly incorporation of some of this thinking is already occurring as evidenced by the newest of the criteria document efforts (e.g., Cu, ammonia, cyanide, selenium). By focusing efforts on those concerns that are most pressing and developing and conducting appropriate research programs to answer focused problems, it will be possible to move toward full implementation of the "new" criteria model in the future.

Charge Question 2.1. Please comment on whether the kinetic toxicity models being considered by EPA are scientifically appropriate for use in deriving water-based criteria.

Kinetic-based modeling is the correct way to move forward in addressing questions about time-variable concentrations and organism responses. An understanding of toxicant mode of action leading to the development of toxicokinetic/toxicodynamic models is the best way to quantify and predict risks to individual organisms and to address questions associated with exposure duration and magnitude. The key to understanding and predicting toxicity is knowing the "dose" of a material delivered to the target site-of-action. These models are common in mammalian toxicology and in many cases form the bases for our ability to predict effects to individuals. Unfortunately much of the basic

knowledge necessary to develop such models for non-mammalian organisms does not exist. In some cases attempts to use such models have been successful, e.g., copper biotic ligand model and the hydrocarbon narcosis model, and have (at least in the case of Cu) been incorporated into the criteria approach. More research should be expended to further our understanding of toxicant mode of action---it would benefit efforts such as cross-species extrapolations, prediction of toxicity based on structure activity relations, evaluation of tissue-based concentrations.

Charge Question 2.2. Please comment on whether the population models being considered by EPA are scientifically appropriate for use in deriving water-based criteria.

The concept of incorporation of population models into the criteria derivation procedure is a positive step forward and attempts to address concerns that have long been at issue with the evaluation of population-based effects. However, the adequacy of the proposed population models for use in criteria development is unproven and will require a substantial research commitment. It remains unclear how non-contaminant related factors such as habitat-loss, species interactions, disease, etc would be incorporated in the evaluation strategy. Density-dependent models assume systems are at carrying-capacity; I doubt that this is a valid assumption for most systems. I also doubt that sufficient information exists for many species to adequately address post-exposure recovery.

Charge Question 2.3. Please comment on whether the proposal for aggregating effects across species being considered by EPA is scientifically appropriate for use in deriving water-based criteria.

The proposed aggregation approach appears to have the potential to be a step-forward over the current Species Sensitivity Distribution (SSD) approach. Although I find the description of the process appealing; additional details about the procedure would be beneficial. I am concerned about how the required level of protection will be defined. The Consultation document suggests that the Criteria Development Committee favors a single percentage approach; I am concerned that this may be an oversimplification for regulatory expedience and will not appropriately consider factors such as resilience and ecosystem function in the process. Adoption of this type of procedure will no doubt necessitate a greater degree of “best professional judgment” and interpretation into the criteria development process. It is also not clear how the procedure will address the issue of small datasets.

Charge Question 2.4 Please comment on whether the framework being considered by EPA for deriving water-based criteria is scientifically appropriate for use in deriving the criteria.

The USEPA is to be commended for attempting to incorporate more “state-of-the-science” considerations into the criteria development process. The approach builds on the back of the previous methods and attempts to address perceived shortcomings in the existing procedure--this can only improve the adequacy and reliability of the criteria in the end. That said, I am concerned that the proposed model is extremely “ambitious” and

may require a substantial long-term research investment on the part of EPA and industry to develop and refine the proposed approach. I would recommend that a “tiered” strategy be developed that would focus research efforts on the most beneficial changes in the short-term while developing longer-term initiatives to address some of the more “intractable” concerns. It may be possible to address some of the initial concerns using extant data with some of the more “data rich” compounds. For example, it may be possible to evaluate population models and aggregation approaches with some of the pesticides for which substantial data exists.

All in all I believe that this is a major step forward in the criteria development approach and that it will ultimately result in improved, more scientifically valid criteria.

1. Overview:

In the mid 1980s, at a symposium on “Pollution and Physiology of Marine Organisms” I asked Charles Stephan (the first author of the 1985 aquatic criteria guidelines) a question: “To what extent has the knowledge of physiological and other sublethal effects of contaminants influenced the setting of water quality criteria?” The answer was “not at all.” Now we are two decades later in the 21st century, and I am sad to see that the many thousands of articles on sublethal effects of contaminants, while given some lip service in the present document, still do not seem to be having much influence on the process. The document on page 6 states “Considerable advancements in the areas of aquatic sciences, aquatic and wildlife toxicology, population modeling, and ecological risk assessment that is relevant to deriving aquatic life criteria has accumulated since 1985.” I fail to see much impact of this in the document. The field of endocrine disruption alone has produced thousands of articles over the past decade showing that concentrations that had previously been considered safe can produce alterations to the endocrine system that can have profound effects. Considering that lethal effects are (happily) seldom seen in natural systems, the continued reliance by EPA on the blunt instrument of LC 50 values is both astonishing and disappointing. It is hard to imagine anything less ecologically realistic than a “standard toxicity test.” Even using LC50 data, it has been shown that when tests have more ecological reality, for example predator cues are present, pesticides become considerably more (up to 46 times more!) lethal to amphibians.

Responses to this comment will be that “there is so much more data available on lethal levels.” The reason for this is that EPA still values such data more than sublethal, more realistic, responses. In the risk assessment paradigm the “effects analysis” needs to incorporate information from the many thousands of studies that have been done on immunological, developmental, genotoxic, neurobehavioral, endocrine, and other sublethal effects. The continued focus on lethal levels has sad implications for the development of water quality criteria, and also has impeded the development of the field of ecotoxicology. While qualified scientists are spending their time doing LC50 tests, the creativity and development of this field is stifled.

The population modeling design, while a good approach, seems to have a limited view of “population effects” in that it recognizes only population size as an endpoint of concern. There is more to population level effects than just population size. If a contaminant causes individuals to grow more slowly and/or not live as long, the population size may remain the same but there will be changes in age and size structure that are also important and are useful and important endpoints of effects.

In considering (p 13) the most sensitive life stages, there is no consideration of the possibility of delayed effects, i.e. exposure at the embryo or larval stages producing effects later on – such effects have been seen with neurobehavioral studies and endocrine disruption. These delayed effects are among the “considerable advancements in the areas of aquatic sciences, aquatic and wildlife toxicology, population modeling, and ecological risk assessment...that has accumulated since 1985.”

Flexibility (p 20) is in theory a good thing and I would support it scientifically. However, I am concerned that in the current political climate (in which science is distorted and misinterpreted for political reasons) pressure will be brought to use the less protective approaches when there are alternate ways of deriving criteria.

The data requirements table on p 21 talks about NOEC and LOEC for characterizing effects on organisms. I want to know if the NOECs and LOECs used by EPA have been lowered for chemicals when new data come in on previously unstudied effects, such as endocrine disruption and other sublethal responses. Many of these effects occur below levels that were previously considered “no effect” levels.

Regarding criteria types, it makes sense to have water-based and tissue-based criteria. (Question: What ever happened to sediment criteria??)

Appendix A, #4 talks about “application of approaches to assess sublethal effects” which reinforces the impression that all the other effects considered are lethal. The “assessment endpoints” on p 28 says “*consider* indirect toxicity and sublethal effects.” Given that we are in the 21st century, I contend that sublethal effects should be the main focus, not this extra additional thing that maybe also should be considered. Also on this table, the physico-chemical factors should include salinity for estuarine environments.

2. Water-Based Criteria

P 4- 1.2 (3) chronic toxicity ought to include a variety of sublethal effects, not mortality.

P 7 The table of taxonomic sensitivity to ammonia does not explain what effects are being measured in the 9 different organisms to obtain the EC50s. If they are comparing growth in one species, fecundity in another, and hatch rate in a third, this is apples and oranges. Unless the same effect was studied in all the different species, the comparisons are meaningless. What effect was studied?

p 9 – How much reduction is an “unacceptable” reduction in a population?

P 12 - The deterministic process model relies on mortality. Here we are again.

P 16 discussing density dependence and independence. A population model should reflect that population growth is a density dependent phenomenon. That’s how populations work. A density-independent model does not relate to the real world, so there is little reason to bother with them. You should do your best with density-dependent models.

Question: But are effects of pollutants density-dependent or independent?

P 24 – recovery time is usually gradual. There is a gradual reduction in the level of the contaminant, and gradual recovery of the population. There are long term field data on oil spill recovery that might be useful in this regard.

P 25 – To repeat, the approach of population impairment focuses only on missing individuals ignoring the possibility of reduced growth and longevity while the population size could remain the same. The population models should be enhanced to cover age and size structure.

3. Tissue Based Criteria

p 7 – The “limitations” of the 1985 guidelines should include their reliance on acute toxicity data and ignoring of sublethal effects.

p 9 top “the toxicological basis for these criteria is driven mostly by measurements of acute lethality.” Here we are again...

p 11 Regarding the Generalized Schematic (also on p 37, 45)– what is the definition of “minimum data requirements?” I assume that means a lot of acute lethality data on a bunch of different taxa. It ought to require a lot of sublethal data.

P 21 “tissue type” indicates that the preponderance of data is whole body concentrations. This is true of small organisms, but it would not be the case for large fish or birds and other wildlife.

P 22 first paragraph says that life history attributes such as generation time “may also be considered” – this is very important and should be considered.

P 24 extrapolating across magnitudes of effect – refers to NOAELs and LOAELs and EC50s. These numbers are not meaningful unless the particular “E” (EFFECT) is described and is consistent in all organisms studied.

P 33 also refers to NOAELs and LOAELs. Have these numbers been revised by EPA to reflect new data on sensitive endpoints – e.g immunological, genotoxic, neurobehavioral, and endocrine etc. effects – in the past 20 years?

P 38 The second bullet refers to reproductive endpoints and says (as an example) that the number of fledglings produced per nesting attempt is a better endpoint than the number of eggs laid. True, but it is important that these fledglings be healthy and not compromised in terms of their immune system, endocrine, and neurobehavioral etc. systems. That’s why it’s important to look at all these things.

P 42 – trophic transfer factors. It is possible that not all the chemical transferred is trophically available. This is the case for metals that may be transferred in granules, which are not biologically available.

P 43 – In the last paragraph, it should be taken into account that a particular species may not be at the same trophic level in all ecosystems. If an invasive species, for example, adds additional link to the food chain, the top predator will be one step higher than in ecosystems without this invader.

P 49 Questions:

1. The rationale and conceptual approach is fine, it just needs to take more things into consideration.
2. Flexibility is a good thing, in theory. Again, I am concerned that in the current political climate, there will be pressure to use the least protective approach.
3. Population modeling needs to consider other things besides the total numbers of the population such as age structure and size structure.

P 50 Questions:

2. Surrogate species have to be used if there are no data on the species of interest. I don’t see any other alternative.

P 58 Appendix – 3. “mortality is by far the most common endpoint measured...” Why is this the case?? Because EPA has required this kind of data and not the more ecologically realistic and useful sublethal effects. If EPA asked for sublethal effects, you would get them.

4. Taxon-Specific Criteria

P 6 – Technical issues 1. Lowering the criteria below the current fifth percentile might be the easiest way to deal with these issues.

2. Yes it is feasible

3. Problem formulation issues: For many endangered species, it is unlikely that excessive contamination is the prime reason for their scarcity. For example, salmon have to deal with clear-cut forests, multiple dams to navigate both going up as adults and going down as juveniles, stream damage from livestock, and overfishing. Given these major multiple risks the added impact of pollution is probably a drop in the bucket. So even if salmonids are especially sensitive, going through this long exercise to reduce numerical criteria will make EPA staff feel that they are doing good, but will probably not help the salmon in any meaningful way. It may not be the best way for EPA staff to spend their time unless it will really help the populations of the species of concern. My guess is that there are some amphibians that are getting scarce because of pesticides, but I don't have the data.

P 8 – 2.3.2.1 interspecies correlation estimates. These are useful ways to compare species. Too bad they are based only on lethal concentrations (yet again).

P 10 2.3.2.7– uncertainly factors – empirically derived safety factors are preferable to arbitrary ones. Of course – but how do you get an empirically derived safety factor? Most of EPA's safety factors are 10, which is totally arbitrary and not empirically derived.

2.4 Guidance – 1. Role of taxon specific criteria to protect specific taxa. We need to find endangered taxa that are endangered because of too much contamination and will be helped through lowering the criteria. I don't know of any, but would like to learn of some.

P 12 technical issues. 1. There is a stress on “scientifically defensible” criteria. If some species are put at risk because of contamination (amphibians with pesticides may be a case), lowered criteria are needed and thus scientifically defensible. If there is no evidence that it is contamination rather than habitat factors (such as with salmon) causing the problems, they won't be scientifically defensible. If species are endangered through loss of habitat and overfishing, lowering the criteria will not help them.

2. Surrogate species are needed. Too bad the ICES are based on only lethal concentrations!

3. Tools to use – it might be easiest overall to just lower the 5th percentile level of protection.

Charge Questions:

1. Scientifically-defensible numerical criteria – You may need to show that pollution is responsible for at least some of the problems that the species is experiencing.
2. Surrogate species have to be used in some cases. Viewing the concern about procedures being “scientifically defensible” how does the 10% safety factor used in so many risk assessments pass muster as “scientifically defensible”??

Final Report Summary from 2003

P 1 emphasizes the importance of incorporating new and emerging science and reflecting the latest scientific knowledge. Yet I don't see it in the new document regarding such sublethal effects as endocrine disruption, genotoxicity, developmental toxicity, neurobehavioral effects etc. where there are decades of work and many thousands of studies on these topics.

P 5 kinetic based modeling – is still relying on survival counts – lethality. Here we are again.

P 7 says a parallel approach “could be applied” to sublethal stresses. The appropriate word is “should” not could.

The writers seem to think that concentrations causing sublethal effects are only slightly lower than lethal levels. I strongly disagree!! Many endocrine effects are seen at levels below what was previously believed to a no-effect level. Furthermore, the focus on “most sensitive life stage” does not consider that exposures during embryonic stages can lead to delayed effects that are seen later in adults.

P 8 recommendations for population modelling – should also consider size and age structure, not just population size.

P 13 – They consider early life stage tests with fishes to be acceptable surrogates for full life cycle tests. However, this is not a good idea since early life stage tests won't tell you if the fish will grow up to have reproductive, immunological, neurobehavioral or other problems.

P 14 – 8a rapid-chronic tests – this is an oxymoron.

8b “Chronic data are expensive and therefore sparse.” There have been thousands of published studies on sublethal chronic effects that don't seem to have made any difference to EPA, even now in the 21st century. What ever happened to “incorporating new and emerging science and reflecting the latest scientific knowledge”?

P 15 8d recommendations – considering the variety of sublethal effects that have been studied by numerous investigators around the world, these papers are being ignored. I don't believe that these effects occur at levels just slightly below lethal levels, except in some very unusual cases.

P 15-16 – physicochemical factors – don't forget salinity for estuarine species

P 20 – Again it is stated that criteria should be as “ecologically relevant” as the state of the science will allow. This advice appears to have been ignored with a focus on lethal levels, which don't occur in nature except in the most dire circumstances.

I repeat: What ever happened to “incorporating new and emerging science and reflecting the latest scientific knowledge”???

P 26 15b– They say don't use acute toxicity when organisms were fed because “standard test methods use unfed animals.” These are the people who ask for ecological relevance???

P 27 15e The group opposes using growth and reproductive data!! This is amazing! Who wrote that they wanted ecological relevance??

15g – tests under unnatural conditions should be excluded. The group is opposed to this recommendation – I am totally befuddled by this.

P 30 – At the end there is a mention of endocrine disruption as a non-traditional endpoint. This is lip service only, since these endpoints have had no impact on the general procedures advocated in the report.



Revising EPA's Aquatic Life Criteria

*Presentation to the Science Advisory Board
September 21, 2005*

*Presented by
Edward Ohanian, Director
Health & Ecological Criteria Division*

*On behalf of
Office of Science and Technology
Office of Water
U.S. Environmental Protection Agency*

1

Ambient Water Quality Criteria

- Section 304(a)(1) of the Clean Water Act requires EPA to develop and publish, and from time to time revise, criteria for water quality accurately reflecting the latest scientific knowledge.

2

Aquatic Life AWQC

- The methodology by which EPA derives AWQC for aquatic life protection was published in 1985:
 - *Guidelines for Deriving Numerical National Water Quality Criteria for the Protection of Aquatic Organisms and Their Uses* (Stephan, Mount, Hansen, Gentile, Chapman, and Brungs 1985)

3

Guidelines Revision

- EPA's Office of Water and Office of Research & Development recommended updating/revising the *Guidelines* and identified issues that should be addressed in revisions.
- EPA's Science Advisory Board concurred with the need to update the *Guidelines* and with the issues EPA identified to address (2002).
- Updating the *Guidelines* is a Priority Action in the *Strategy for Water Quality Standards and Criteria* (2003).

4

The Aquatic Life Criteria Guidelines Committee

- Interagency Committee
- EPA Offices:
 - Water (OW)
 - Research & Development (ORD)
 - Pesticide Programs (OPP)
 - Solid Waste & Emergency Response (OSWER)
 - Regional Offices (3,4,5,8,9)
- Other Federal Partners:
 - U.S. Fish & Wildlife Service
 - NOAA - Fisheries

5

The Aquatic Life Criteria Guidelines Revisions – Process

- Subcommittees → Committee
- Committee → OW
- OW → Required Reviews/Input :
 - Internal Agency review
 - EPA Science Advisory Board review
 - External peer review
 - Public/Stakeholder review & input
- OW issues final guidance

6



Overview of Proposed Revisions to EPA's Aquatic Life Criteria Guidelines

*Presentation to the Science Advisory Board
September 21, 2005*

*Presented by
Tala Henry, Office of Water
U.S. Environmental Protection Agency*

*On behalf of
The Aquatic Life Criteria Guidelines Committee* ⁷

Update/Revision History

- Since 1985 -incorporated latest science on criteria-specific basis:
 - Development of Wildlife Criteria Methodology in Great Lakes Water Quality Initiative (1995)
 - Use of concentration-response modeling in Ammonia Criteria Update (1999)
 - Use of population modeling in: Aquatic Life Water Quality Criteria for Dissolved Oxygen (Saltwater): Cape Cod to Cape Hatteras (2000)
 - Use of biotic ligand model in: Draft Update of Ambient Water Quality Criteria for Copper (draft 2003)
 - Development of tissue-based criteria for Selenium Aquatic Life Criteria (2004)

8

The Aquatic Life Criteria Guidelines Committee - Charge

- Revise and update the methodology for developing aquatic life criteria.
- The revised *Guidelines* will:
 - Incorporate latest scientific Approaches into AWQC derivation
 - Be less prescriptive
 - Provide flexibility for incorporating a variety of risk-based approaches, methods, and models and enhance ability to make site-specific adjustments

9

Key Issues

- Consider wider range of organisms & effects
- Improve recommendations on averaging period & allowable frequency of criteria exceedances
- Integrate organismal effects and extrapolate to population level
- Better assess of bioaccumulative chemicals
- Better characterize risk and quantify uncertainty

10

The Criteria Guidelines Committee Organization

- Committee
 - Tala Henry, Chair

- Three Subcommittees
 - Water-Based Criteria Subcommittee
 - Chuck Stephan, Chair
 - Tissue-Based Criteria Subcommittee
 - Aquatic Life: Keith Sappington, Chair
 - Wildlife: Rick Bennett, Chair
 - Taxon-Specific Criteria Subcommittee
 - Kellie Kubena, Chair

11

Types of Criteria

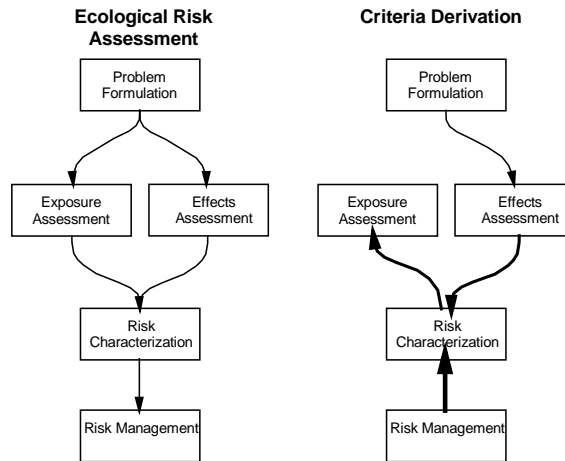
- **Water-based Criteria:**
 - for chemicals for which water concentration is a reasonable predictor of effects expected under natural exposure conditions
 - will be derived and expressed as water concentrations

- **Tissue-based Criteria:**
 - for chemicals for which water concentration is not a reasonable predictor of effects expected under natural exposure conditions, such as:
 - chemicals that bioaccumulate and/or biomagnify;
 - chemicals for which diet is an important exposure pathway), and
 - will be derived and expressed as either tissue concentrations and/or as water concentrations (by incorporating appropriate tissue-to-water translation procedures).

- **Taxon-specific Criteria:**
 - Modification of the aforementioned criteria types to provide appropriate levels of protection for specific taxa

12

Ecological Risk Assessment and Criteria Development Share Key Components



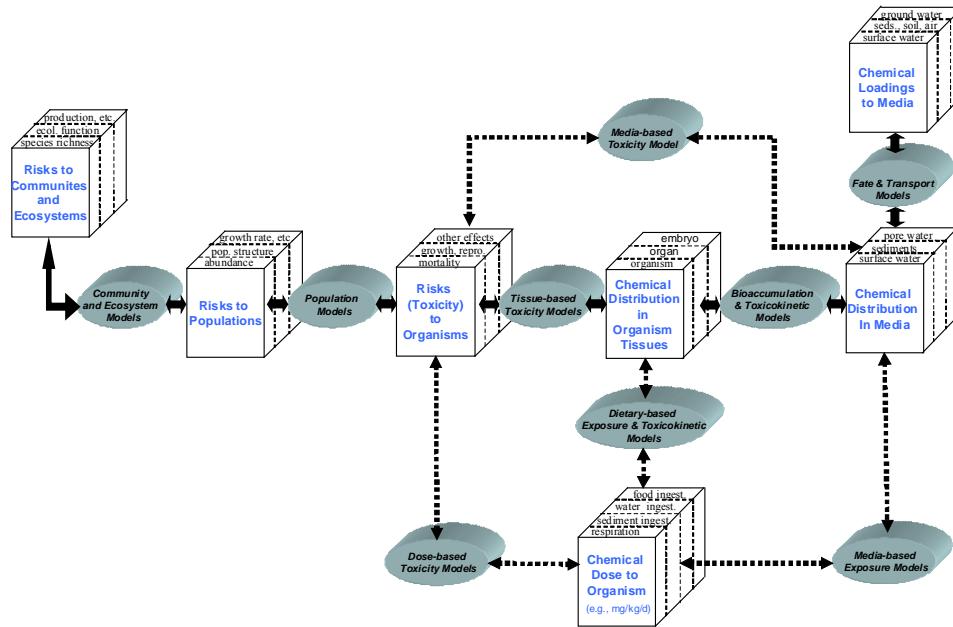
13

Problem Formulation

- “National-level” Problem Formulation
 - National CWA Goals
 - Generic Assessment Endpoints
 - General Conceptual Model
- Criteria-specific Problem Formulation
 - Refined Conceptual Model - determine criteria type(s)
 - Chemical- and/or Species-specific Assessment Endpoints
 - Measure of effect appropriate for chemical mode of action and most vulnerable species

14

Generic Conceptual Model for Criteria Derivation



Analysis

- Criteria Guidelines are the “National-level” Analysis Plan
 - Provide national consistency, but flexibility in presenting multiple approaches and guidance
- Criterion-specific problem formulation will result in a criterion-specific analysis plan
- Criteria Derivation is the Analysis Phase

Risk Characterization

- Criteria-type and criterion specific
- Adherence to Agency guidance
 - Risk Characterization Policy (Browner, 1995)
 - Risk Characterization Handbook (2000)
- Focus of future efforts:
 - risk management discussions & decisions

16

Charge Question 1

- Please comment on the use of the *Guidelines for Ecological Risk Assessment* as an essential and relevant organizing framework for development of science-based criteria for the protection of aquatic life and aquatic-dependent wildlife?
- Does the SAB have any specific recommendations on how to improve or clarify the generic conceptual framework diagram?

16

Charge Question 2

- Please comment on whether the proposed criteria types and the scientific focus for each criteria type are logical and scientifically valid for developing a holistic and integrated criteria framework.



Proposed Revisions to EPA's Aquatic Life Criteria Guidelines: Water-based Criteria

*Presentation to the Science Advisory Board
September 21, 2005*

*Presented by
Charles Delos, Office of Water
U.S. Environmental Protection Agency*

*On behalf of the
The Water-based Criteria Subcommittee*

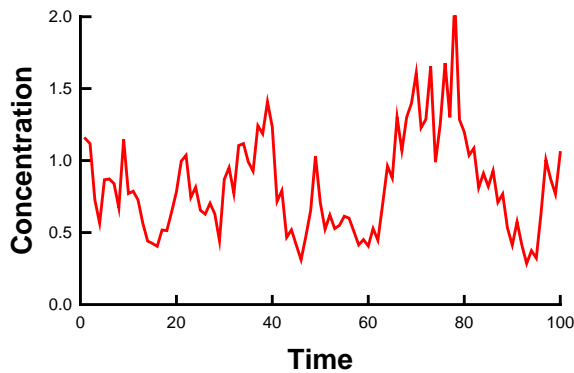
Consider the Sensitivity of Tested Species

Genus	Chronic EC20 (mg/L)
Ceriodaphnia	16.10
Daphnia	12.30
Ictalurus	8.84
Catostomus	4.79
Micropterus	4.56
Pimephales	3.09
Lepomis	2.85
Muscullium	2.26
Hyaella	1.45

2

New Proposal

Address Time-Variable Concentrations Throughout the Criteria Derivation Process

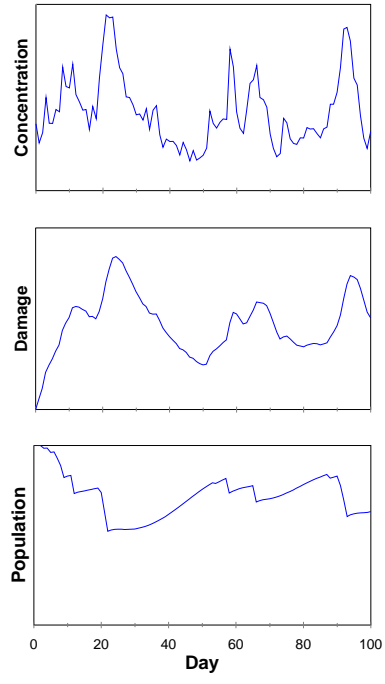


3

For Each Species, Apply Two Models:

- Kinetic toxicity model to translate from lab test exposures to continuously variable concentrations.
- Life-stage structured population model to account for:
 - Population reduction from effects on survival and reproduction.
 - Rate of recovery after population loss.

4

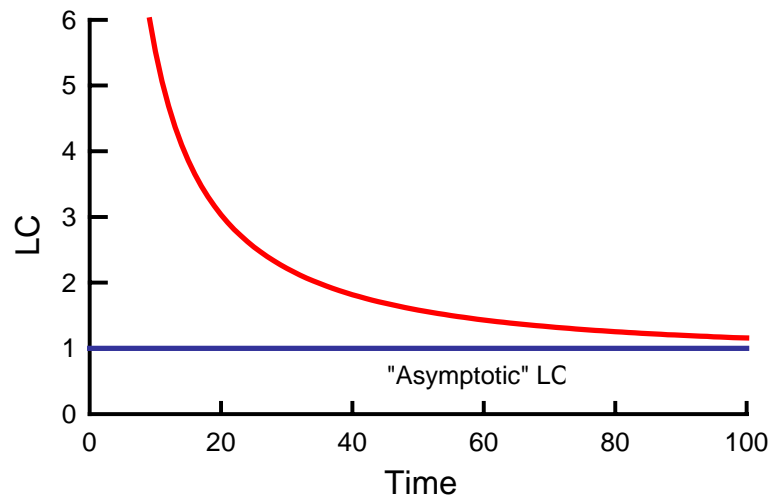


5

Kinetic Toxicity Model

6

Effect of Time on Lethal Concentrations



7

Kinetic Toxicity Model

- Chemical (or damage) accumulates by one-compartment, first-order kinetics:

$$A(t) = \frac{k_U}{k_E} \int_0^t (k_E \cdot e^{k_E(x-t)} \cdot C(x)) dx = BCF \cdot \bar{C}(t)$$

($\bar{C}(t)$ is a weighted running average water concentration, the weighting factor exponentially decaying backward in time according to k_E .)

- Multiple compartments (additional degrees of freedom) can be used as data warrant.

8

Kinetic Toxicity Model

- The influence of accumulation on survival can be evaluated from two perspectives:

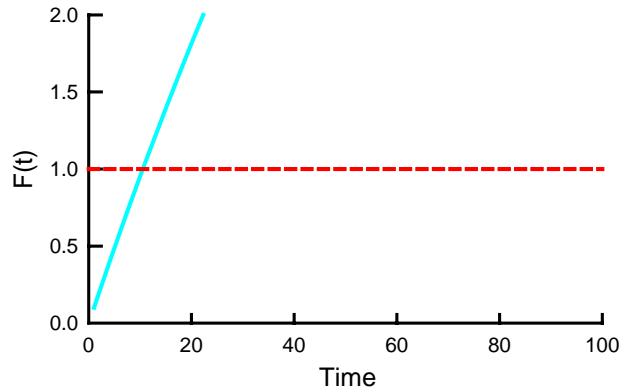
Deterministic Process Model

Stochastic Process Model

9

Deterministic Process Model

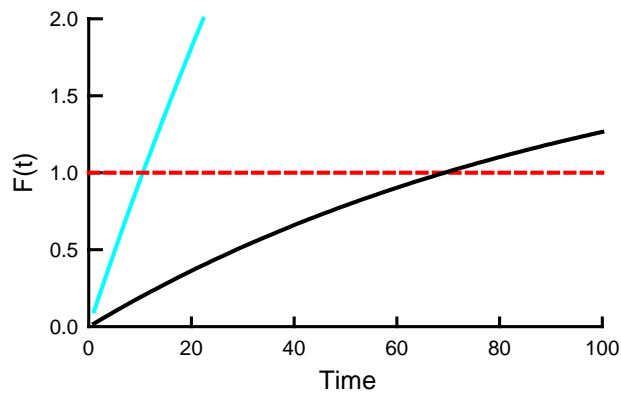
- An organism will die when a lethal accumulation is exceeded. $F(t)$ = fraction of lethal accumulation.



10

Deterministic Process Model

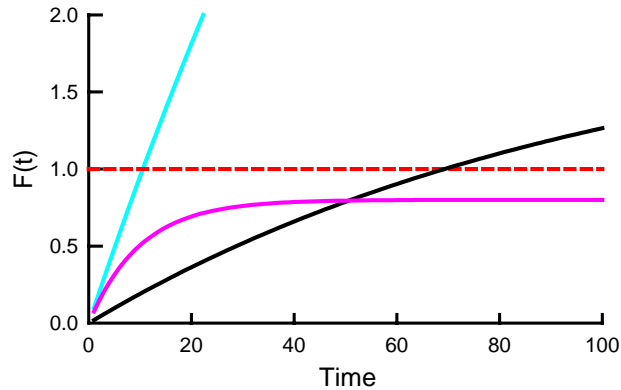
- An organism will die when a lethal accumulation is exceeded. $F(t)$ = fraction of lethal accumulation.



11

Deterministic Process Model

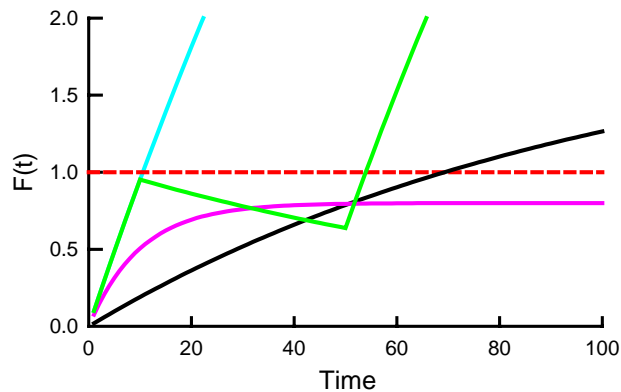
- An organism will die when a lethal accumulation is exceeded. $F(t)$ = fraction of lethal accumulation.



12

Deterministic Process Model

- An organism will die when a lethal accumulation is exceeded. $F(t)$ = fraction of lethal accumulation.



13

Deterministic Process Model

- Lethal condition can be expressed in terms of either accumulation or water concentration:

$$F(t) = \frac{A(t)}{A_0} = \frac{\bar{C}(t)}{C_0}$$

(where A_0 is the lethal accumulation and C_0 is the threshold lethal water concentration for an individual.)

- Model parameters (k_E , C_0) are for individual organisms; groups of individuals are described with distributions of parameters.

14

Stochastic Process Model

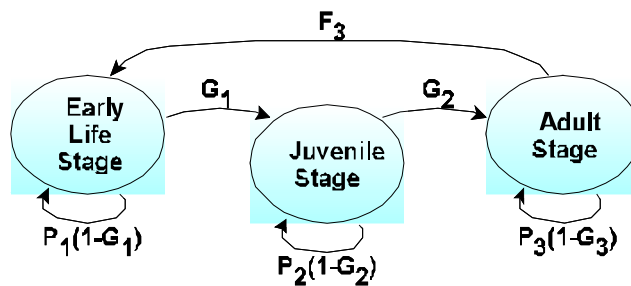
- An individual might die after the threshold lethal accumulation is exceeded.
- The probability of death per unit time is proportional to how much the accumulation exceeds the threshold lethal accumulation.
- All individuals have the same model parameter values.

15

Population Model

16

Life-Stage Structured Population Model



Stage model diagram with survival probability P , graduation probability G , and fecundity F .

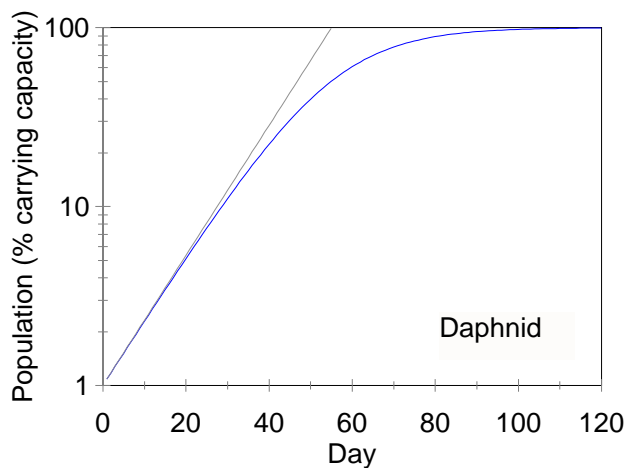
17

Two options for formulating life-stage structured population model

- Density-independent
 - If survival or reproduction is not formulated as a density-dependent process, then the population can grow exponentially without limit.
- Density-dependent
 - Density-dependent reduction in survival or reproduction causes the population to level off at a carrying capacity.

18

Density-Independent versus Density-Dependent Growth



19

Population Model Parameters

Parameters for both density-independent and density-dependent models:

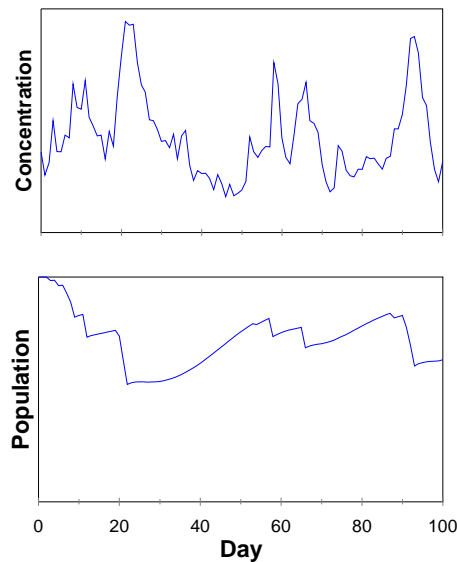
- Stage durations
- Background death or survival rates
- Fecundity of adult stage(s)
- (Time-variability of parameters)

Density-dependent parameters (applied to survival or reproduction):

- Mathematical function that describes density dependency
- Parameter value for each life stage

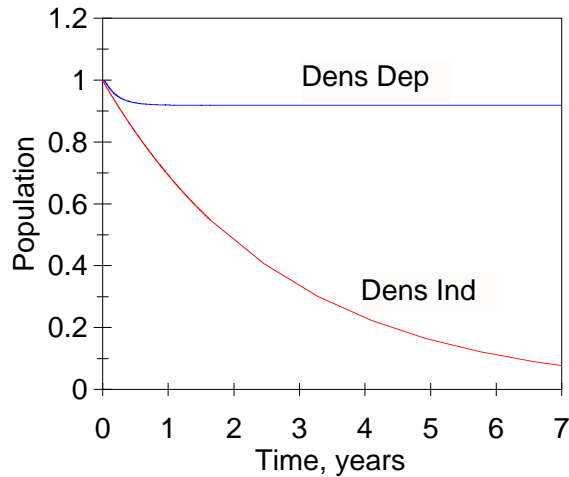
20

Population Assessment: Density-Dependent Model



21

Behavior for Constant Exposure at EC06: Density-Independent versus Density-Dependent



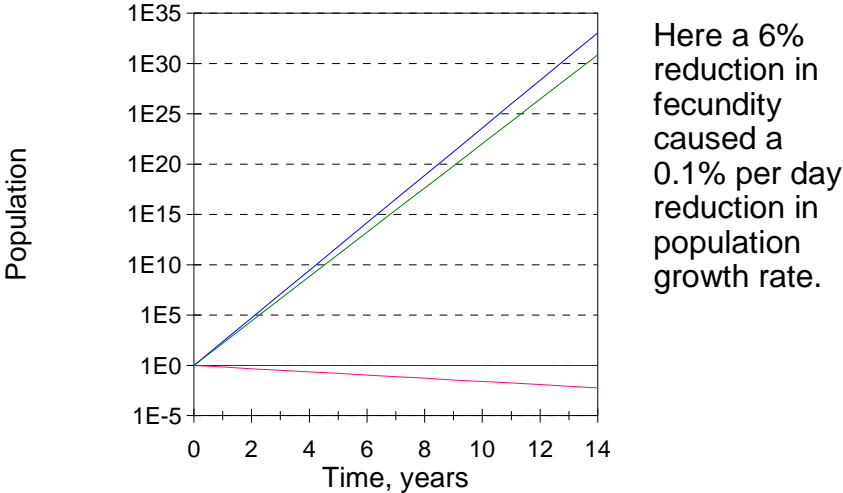
23

Density-Independent Model versus Density-Dependent Model

- Density-independent model would not be used in the same way as a density-dependent model.
- **Density-independent model** might be used:
 - To compare growth rates.
 - To assess probability of falling below some specified density: **Population Viability Analysis**.

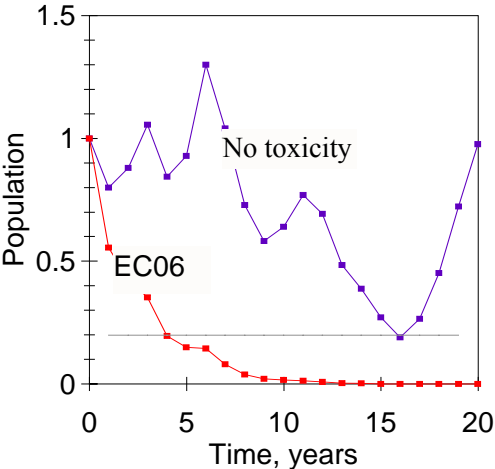
24

Growth Rates in Density-Independent Model



25

Illustration of Density-Independent Population Viability Analysis



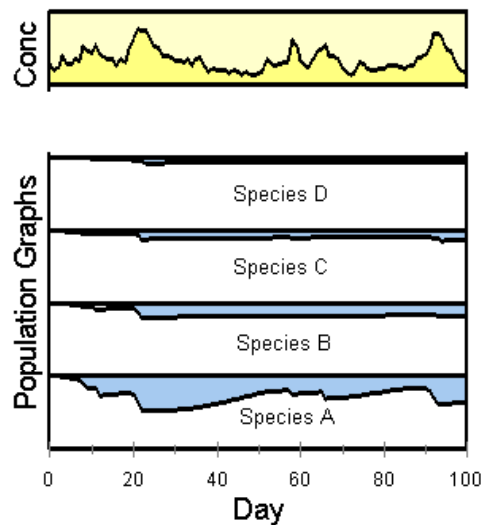
26

Density-Independent Model versus Density-Dependent Model

- **Density-dependent model** can be used to assess percentage of individuals missing due to toxicity.
- Pros
 - Accounts for recovery time.
 - Assessment endpoint is easy to understand.
- Con
 - Model parameters more difficult to determine.

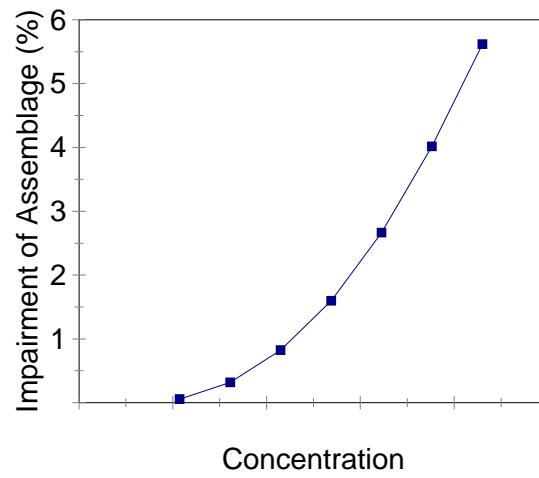
27

Obtaining Criterion that Protects One Assemblage - Using Density-Dependent Model -



28

% Impairment v. Concentration



29

Question on Kinetic Toxicity Model

- Are kinetic toxicity models scientifically appropriate for use in deriving criteria?

30

Questions on Population Model

- Is our preference for attempting use a density-dependent approach reasonable?
- Are there concerns or cautions about obtaining parameter values for population models?
 - Density-dependent
 - Density-independent

31

Question on the Proposal as a Whole

- Is our proposed general structure reasonable?
 - Linking a toxicity model to a population model.
 - Examining the impacts on individual species.
 - Aggregating these impacts into a risk metric for an assemblage of species.



Proposed Revisions to EPA’s Aquatic Life Criteria Guidelines: Tissue-based Criteria for “Bioaccumulative” Chemicals

*Presentation to the Science Advisory Board
September 21, 2005*

*Presented by
Keith Sappington, Office of Research & Development
Richard Bennett, Office of Research & Development
U.S. Environmental Protection Agency*

*On behalf of the
The Tissue-based Criteria Subcommittee*

1

Problem Statement

- EPA’s *1985 Guidelines* do not comprehensively address ecological risks from “bioaccumulative” chemicals.
 - Bioaccumulation methods are limited in scope
 - Toxicological guidance is relatively sparse

2

Proposed Solution Being Considered

- “Tissue-based” approach for aquatic life and aquatic-dependent wildlife
- **Aquatic life:** based on chemical concentrations in tissue, e.g.,
 - *2,3,7,8-TCDD*: Dioxin interim report (EPA et al., 1993) and Steevens et al., 2005
 - *Selenium*: draft aquatic life criterion (EPA 2004)
- **Wildlife:** based on chemical concentrations in tissue or aquatic diet, e.g.,
 - Great Lakes Wildlife Criteria, EPA Mercury Study Report to Congress

3

Current Challenges with a Tissue-based Approach

- Available tissue-based toxicity data appear limited in several ways:
 - Relatively few species per chemical
 - Mostly whole-body measurements

Sublethal endpoints (reproduction, growth) are relatively sparse

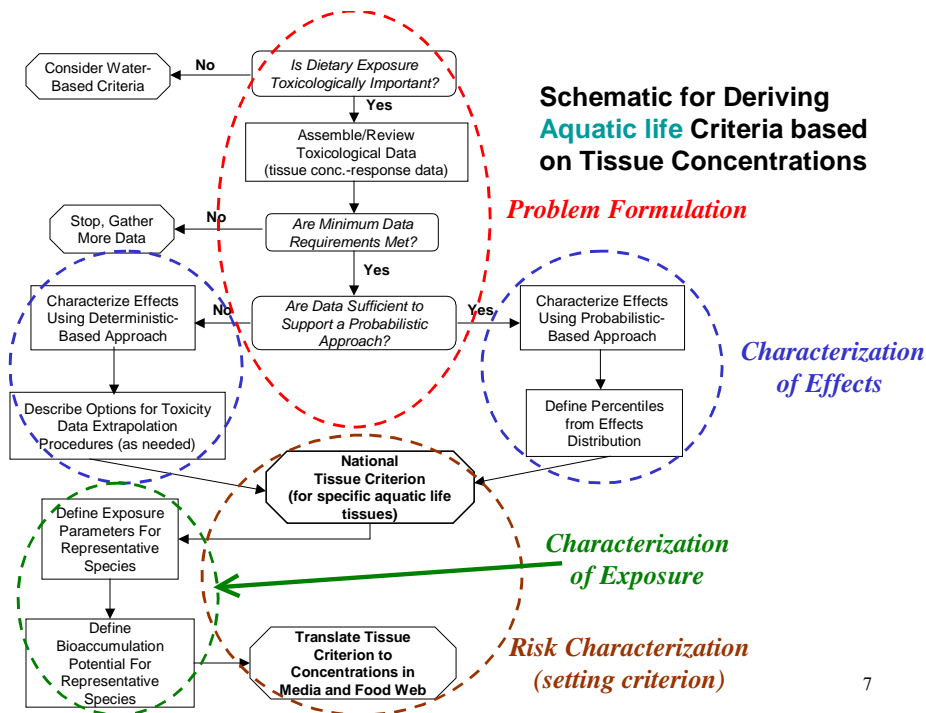
Heterogeneity in test designs

5

Primary Components

- Procedures for deriving a **national tissue criterion**
- Procedures for **translating** a national tissue criterion into concentrations in **ambient media** and components of the **aquatic food web**

6



7

Deriving a Tissue Criterion for Aquatic Life

Characterization of Effects

- Generally, tissue-based toxicity data indicative of chronic exposures
- Three aquatic assemblages
 - (vertebrates, invertebrates, plants)
- Survival, reproduction, growth, other endpoints with strong linkage to population level effects
- Toxicity data extrapolations
 - inter-tissue, inter-species, magnitude of effect, exposure duration
- Population modeling being explored

9

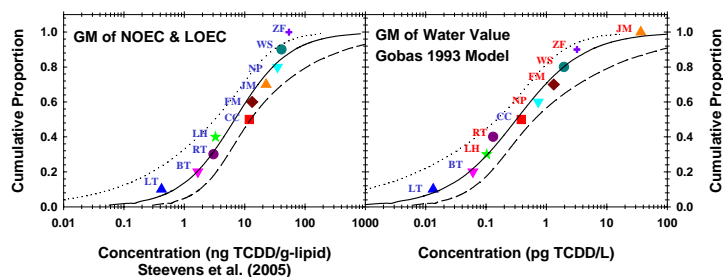
Issues with Deriving a Tissue Criterion for Aquatic Life

- How to conduct extrapolations with tissue-based toxicity data, e.g.,
 - between tissues (empirical, PBTK-based methods)
 - between tested and untested species
 - between exposure durations (if relevant)
- Feasibility and utility of population modeling

11

Translating Tissue Criteria into Concentrations in Media/Food Web

- Issues:
 - Ambiguity in species exposure potential at tissue criterion
 - Need to address discontinuity between species-specific exposure potential and inherent (tissue-based) toxicity



13

Translating Tissue Criteria into Concentrations in Media/Food Web

- Proposal:
 - Use “representative species” coupled with methods for estimating species-specific bioaccumulation potential
 - Representative species could be defined for a range of exposure potentials within an assemblage (e.g., different feeding guilds, habitats, etc.)
 - Representative species could be defined on a site-specific basis

14

Translating Tissue Criteria into Concentrations in Media/Food Web

- Define bioaccumulation potential for each representative species
- Considering a bioaccumulation framework similar to EPA’s human health criteria methodology (EPA 2000, 2003)
 - Field-, lab- and model-derived estimates of bioaccumulation (BAF)
 - lipid and organic carbon normalization (nonionic organics)
 - fugacity-based food web model used (Gobas, 1993) when chemical metabolism is considered negligible
 - Steady-state assumptions, unless temporal variability likely to be important

15

Translating Tissue Criteria into Concentrations in Media/Food Web

- Site-specific estimates of bioaccumulation would be encouraged
- Nationally-representative parameter values might be developed for use when site-specific data are lacking.
- Output: multiple “translated criteria” for each representative sp.

Translated Criterion Concentration	Aquatic Vertebrate Assemblage		
	Representative Sp. A (piscivore)	Representative Sp. B (benthic carnivore)	Representative Sp. C (herbivore)
Water	✓	✓	✓
Sediment	✓	✓	✓
Algae/Macrophytes	✓	✓	✓
Zooplankton	✓	✓	
Macroinvertebrates	✓	✓	
Forage fish	✓		

16

Process for Developing Tissue-based Wildlife Criteria

- Conceptually, the process for tissue-based criteria is the same for aquatic life and aquatic-dependent wildlife
- However, the technical procedures may differ reflecting:
 - differences in the toxicity data available,
 - differences in exposure pathways, and
 - differences in life history

18

Process for Developing Tissue-based Wildlife Criteria

- Historically, wildlife criteria have been based on toxicity data from long-term feeding studies, with dietary concentrations converted to daily ingested doses
- Where it is appropriate and data are available, wildlife criteria also may be based on chemical concentrations in animal tissues (e.g., eggs, liver, brain)

19

Process for Developing Tissue-based Wildlife Criteria

- Where data are available, probabilistic methods should be used to more explicitly address natural variability and uncertainty
 - For example, where toxicity data exist for many species, a species sensitivity distribution may provide a better estimate for a more sensitive species than the use of uncertainty factors
- When data are limited, deterministic methods are more appropriate

20

Focus on National-level Criteria

- National-level criteria would provide an analysis of all available toxicity data and background on the parameter estimates used for representative species.
- May be adopted by State, Tribal, or local agencies **or** may be modified at state or local scales *if* sufficient additional information is available to improve the characterization of risk while maintaining the intended level of protection

21

GLI Approach for Wildlife Values

$$WV \text{ (mg/L)} = \frac{TD \text{ (mg/kg bw/d)} * (1/(UF_A * UF_S * UF_L)) * BW \text{ (kg)}}{W \text{ (L/d)} + \sum [FC_i \text{ (kg food/d)} * BAF_i \text{ (L/kg)}]}$$

where:

WV = wildlife value expressed as the chemical concentration in **water** for each representative species,

TD = test dose expressed as daily dietary dose from selected study,

UF = uncertainty factors for interspecies variation (UF_A),

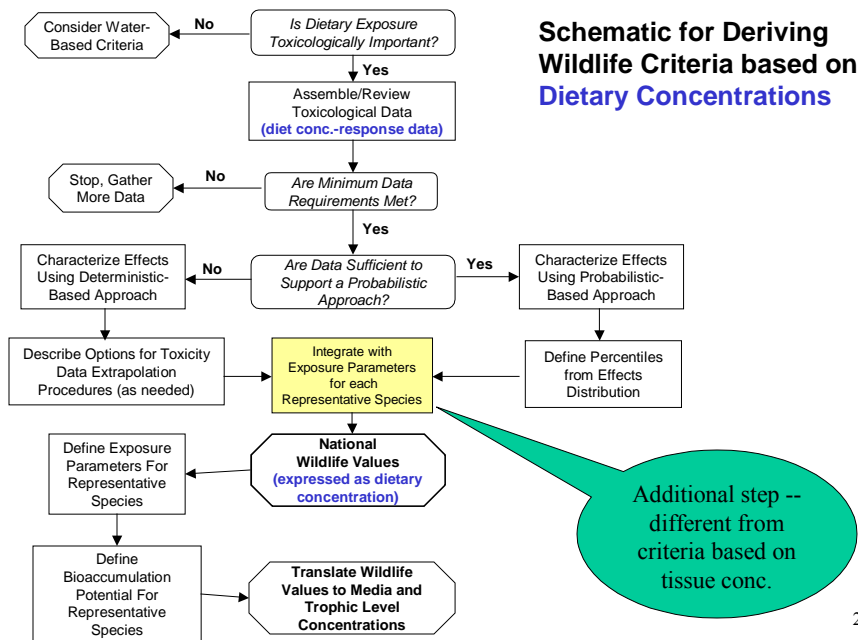
subchronic to chronic (UF_S), and LOAEL to NOAEL (UF_L)

BW = body weight of species of concern

W = amount of daily water consumption

FC_i = amount of daily food consumption from the i^{th} trophic level

BAF_i = bioaccumulation factor for the i^{th} trophic level 22

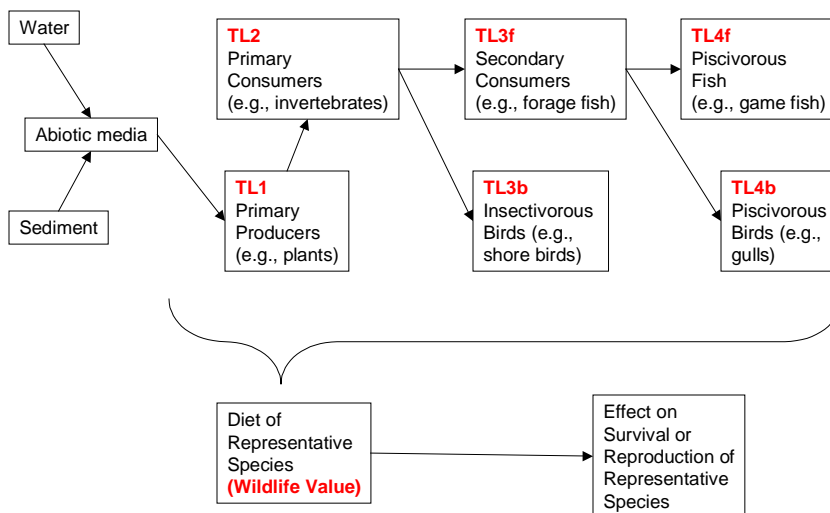


Translating Wildlife Values

- Wildlife values based on whole diet need to be translated into corresponding concentrations in each trophic level
- Translation based on trophic transfer factors and diet composition (% from each TL)
- Provides basis for developing criteria and a common currency for comparing among representative species

26

Relationship between Wildlife Values and Aquatic Foodweb



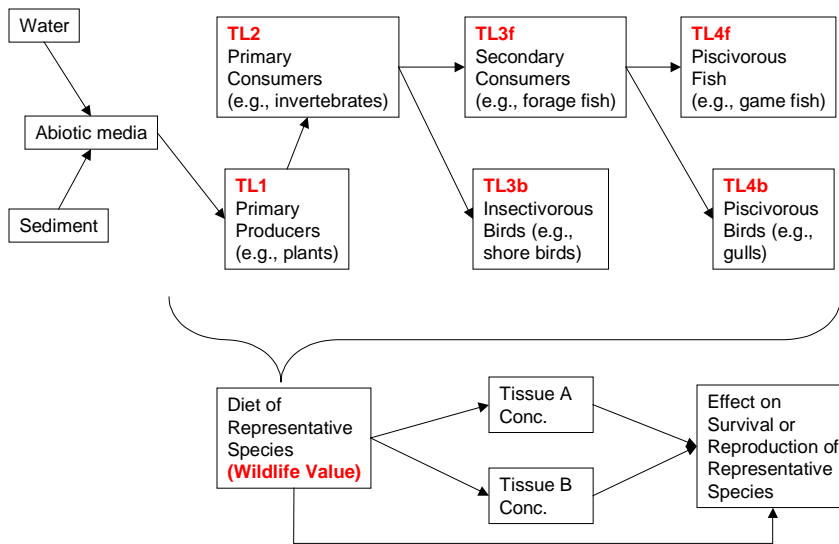
27

Example of Translating Wildlife Values to Corresponding Concentrations in Various Trophic Levels

	Bald eagle	Osprey	Peregrine falcon
Adult body wt (kg)	4.6	1.48	0.78
Total FIR (kg/day)	0.504	0.294	0.125
% TL3 fish	73.6	90	
% TL4 fish	18.4	10	
% TL4 birds	5.6		35.1
% Non-aquatic food	2.4		64.9
TD = 0.078 mg/kg/day; UF _A = 1; UF _S = 1; UF _L = 3			
Wildlife Value (mg/kg diet)	0.237	0.131	0.163
Trophic Transfer Factors: TL3f to TL4f = 4.25 and TL3f to TL4b = 10			
Conc. in TL3 fish (mg/kg)	0.11	0.10	0.05
Conc. in TL4 fish (mg/kg)	0.49	0.42	0.20
Conc. in TL4 birds (mg/kg)	1.14	0.99	0.46

28

Relationship between Tissue Concentrations and Aquatic Foodweb



30

Wildlife Criteria - Summary of Issues

- May be based on chemical concentration in wildlife tissues or their diet
- May be calculated using deterministic or probabilistic methods, depending on data availability
- Currently focused on national-level approach as the basis for refinement at smaller scales when additional information is available
- Methods being developed to translate wildlife values (expressed as diet or tissue) into corresponding concentrations in aquatic food web

31

Charge Question 2

- **Considering the strengths and limitations of the more flexible approach used to derive tissue-based criteria, please comment on the rationale and preference for allowing flexibility in the procedures used.**

33

Charge Question 3

- **Please comment on the rationale used by the Tissue-based Criteria Subcommittee for determining if/when to use population modeling in the development of Tissue-Based Criteria.**



Proposed Revisions to EPA's Aquatic Life Criteria Guidelines: Taxon-specific Criteria

*Presentation to the Science Advisory Board
September 21, 2005*

*Presented by
Brian Thompson, Region 5
U.S. Environmental Protection Agency*

***Tom Augspurger, Fish and Wildlife Service**
U.S. Department of the Interior*

*On behalf of the
The Taxon-specific Criteria Subcommittee*

1

Background

- In addition to national general criteria, natural resource and risk managers may want to ensure protection of “special status” taxa:
 - species or genera **known to be sensitive** to a pollutant (potentially under-protected by the national general aquatic life criteria for that pollutant)
 - taxa that a risk evaluation indicates **may be sensitive and which have a designated special status**:
 - commercial, recreational, cultural, or ecological importance to a Tribe, State or Territory
 - Federally-listed threatened and endangered species

4

Purpose of Taxon Specific Criteria

- Companion national criteria recommendations to provide for the protection of special status taxa as designated by the ESA, State, Territory, or Tribe
 - For use by natural resource and risk managers depending on the level of protection they seek to implement
 - Facilitate State standards development
 - Facilitate Endangered Species Act consultation

6

Purpose of Taxon Specific Criteria

- Companion national criteria recommendations to provide for the protection of special status taxa as designated by the ESA, State, Territory, or Tribe
 - For use by natural resource and risk managers depending on the level of protection they seek to implement
 - Facilitate State standards development
 - Facilitate Endangered Species Act consultation

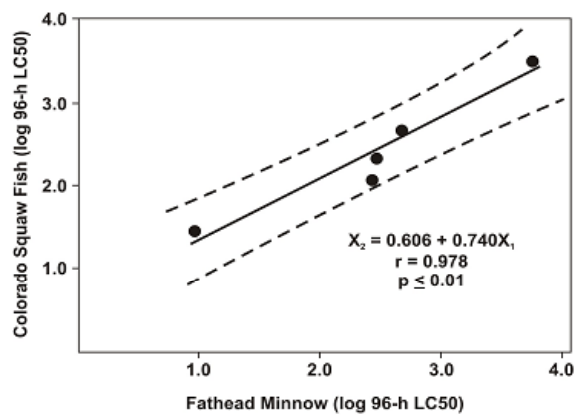
6

Interspecies Correlation Estimates

- Relative sensitivity to other species
 - Estimates made across contaminants
- EPA model for making interspecies correlation estimates
 - Provides strength of relationship
 - Produces a line describing estimate and confidence intervals
 - Estimates made through family
 - Minimum of five chemicals

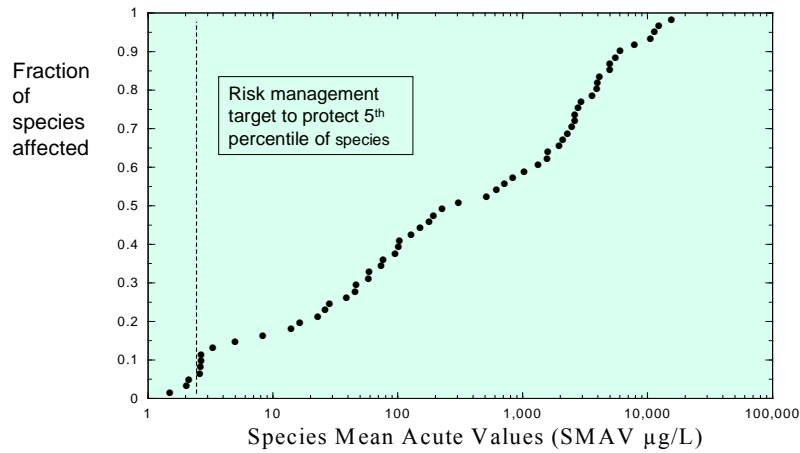
13

Interspecies Correlation Estimation



14

Hypothetical Species Sensitivity Distribution



16

Charge Question 2

- Are the data evaluation/extrapolation tools adequate to develop defensible numeric criteria when there is an absence of toxicological data?
- Are there improvements to the tools that would provide more defensible numeric criteria?
- What other tools are available to provide more defensible criteria?

21

Appendix E – Charge Questions to the Panel

Proposed Revisions to the Aquatic Life Criteria Guidelines

Charge to the Panel

Background:

Since the early 1980's, EPA has been developing ambient water quality criteria to protect aquatic organisms from chemical specific pollutants under Section 304(a) of the Clean Water Act. The criteria provide guidance to states and tribes for adopting water quality standards, which provide a basis for controlling discharges or releases of pollutants. The majority of EPA's aquatic life criteria have been derived based on 1985 *Guidelines for Deriving Numerical National Aquatic Life Criteria for Protection of Aquatic Organisms and Their Uses* (hereafter, referred to as the *1985 Guidelines*; Stephan et al. 1985).

The Agency has recognized the need to update the *1985 Guidelines* for a number of years. Several meetings have taken place since 1985 to discuss additional or complex issues that the *1985 Guidelines* do not currently address. In 1990, EPA convened a workshop, *Workshop on Recommendations for Revising the National Water Quality Criteria Guidelines*. Subsequently, EPA formed the Aquatic Life Criteria Guidelines Committee (hereafter, referred to as the EPA Committee), to guide the work of revising the *1985 Guidelines*. The EPA Committee convened a number of times from 1991-1995 to discuss further approaches for revising the *1985 Guidelines*. The EPA Committee also consulted with the Science Advisory Board (SAB) on several occasions (June 1993; October 1993; April 1994; February 2003) during this time. Proposed approaches for revising the *1985 Guidelines* have also been presented in a number of public and regulatory science forums (e.g. 1994 Water Environment Federation (WEF) Meeting; 2004 Federal-State Toxicology and Risk Assessment Committee (FSTRAC); 2004 Society of Environmental Toxicology and Chemistry Meeting (SETAC); 2005 EPA Science Forum; 2005 Wildlife Society Meeting). The discussions, issues and summaries of the EPA Committee meetings, SAB consultations and other forums concerning *1985 Guidelines* revisions have resulted in a compilation of the major issues, perspectives and suggested directions that should be addressed in proceeding to revise the *1985 Guidelines*. The report, *Summary of Proposed Revisions to the Aquatic Life Criteria Guidelines* (2003), summarizes the discussions, meetings and issues of the Guidelines Committee from 1985 to 1995.

In 2003, the Office of Water reconvened an expanded Committee as the technical body to perform the work of updating *1985 Guidelines*. In addition to the EPA Committee, three subcommittees, namely the *Media-based Criteria*, the *Residue-based Criteria*, and the *Taxon-specific Criteria*, have been formed to address specific technical issues involved in deriving water-based criteria, tissue-based criteria and criteria for the protection of special status species, respectively. These committees include scientists from EPA's Office of Water, Office of Research and Development, Office of Pesticides

Programs and several EPA Regional Offices (4, 5 & 9) and from the U.S. Fish & Wildlife Service and NOAA-Fisheries. During the past two years, the EPA Committee and its affiliated subcommittees have developed approaches and discussed methods for revising the *1985 Guidelines*. The EPA Committee is again consulting the SAB to obtain advice on this framework.

EPA has previously consulted with the SAB regarding development of methodologies for deriving ambient water quality criteria. Feedback from previous Consultations with the Science Advisory Board included the recommendation that the criteria program be guided by EPA's Framework for Ecological Risk Assessment (subsequently developed in more detail as the Guidelines for Ecological Risk Assessment). To date, the EPA Committee has organized its deliberations and development of the conceptual framework for future aquatic life criteria using this EPA guidance. Historically, EPA's ambient water quality criteria for the protection of aquatic life have focused primarily on aquatic organisms and have been expressed as water concentrations. Several limitations are associated with this approach. The SAB has previously concurred with EPA's assessment of the need to revise and also the areas of focus for revising the 1985 Guidelines. The SAB concurred that the 1985 Guidelines revisions should address the following key areas: 1) the use of toxicokinetic modeling to quantify organism responses to time-variable chemical exposures as a basis for improving current recommendations on the frequency and duration of criteria exceedences; 2) the formulation of residue-based criteria that account for multiple exposure routes (e.g. dietary vs. water), bioaccumulation potential and tissue-residue based toxicity relationships to better assess toxicity of bioaccumulative chemicals; 3) the development of methods to better consider assessment endpoints, both a wider variety of ecological entities (e.g. plants and aquatic-dependent wildlife) and attributes (e.g., endpoints other than survival, growth and reproduction, if warranted); 4) the application of population modeling for integrating effects on organism survival, growth, and reproduction and extrapolating these effects to the population level; 5) the development of approaches for characterizing risk to assemblages of aquatic species; and 6) quantifying the uncertainty associated with ambient water quality criteria (AWQC).

Review Material

The EPA Committee has prepared four papers. One paper provides a very brief overview of the existing 1985 Guidelines, focused on highlighting approaches and methods that EPA and others have suggested be improved in the on-going revisions effort. The overview paper also describes general principles and guidance the EPA Committee is using to organize its work and to ensure that the revised Guidelines will provide a logical and scientifically valid approach for developing a holistic and integrated framework for developing ambient water quality criteria in the future. The other three papers describe the approaches and methods being proposed for developing water-based criteria, tissue-based criteria and taxon-specific criteria. The SAB has also been provided a copy of the report, Summary of Proposed Revisions to the Aquatic Life Criteria Guidelines (2003) and the Guidelines for Ecological Risk Assessment referenced above.

Charge to the SAB Panel The EPA Committee is consulting with the SAB to obtain advice on a framework for revising the Guidelines for Deriving Water Quality Criteria for the Protection of Aquatic Life and Aquatic-Dependent Wildlife. EPA seeks comment on: 1) the scope of the proposed framework for revising the 1985 Guidelines, and 2) the scientific validity and appropriateness of proposed approaches for developing water-based, tissue based, and taxon-specific water quality criteria. Specific Charge Questions

1. Scope of the Proposed Framework for Revising the Aquatic Life Water Quality Criteria

The EPA Committee has used the Guidelines for Ecological Risk Assessment as an organizing framework for development of aquatic life water quality criteria. The EPA Committee has developed a generic conceptual model to illustrate the inter-relationships and integration of the various types of criteria and to use as the organizing blueprint from which the methodologies for specific types of criteria are further developed. Given that the scope of future criteria is envisioned to be broader than for the 1985 Guidelines (e.g., better consideration of bioaccumulation, plants, aquatic-dependent wildlife, special status organisms) the EPA Committee has also discussed the strength and necessity of performing a robust Problem Formulation phase as part of the development process for every future ambient water quality criteria (i.e., to identify data gaps or uncertainties that will be inherent in specific criteria; development of conceptual model to assess and inform which types of criteria are necessary and/or feasible; to select assessment endpoints and measures of effect that are appropriate for the stressor of concern).

Charge Question 1.1. Please comment on the use of the Guidelines for Ecological Risk Assessment as an essential and relevant organizing framework for development of science-based criteria for the protection of aquatic life and aquatic-dependent wildlife. Does the SAB have any specific recommendations on how to improve or clarify the generic conceptual framework diagram?

Several types of criteria are currently envisioned by the EPA Committee, namely, Water-Based Criteria, Tissue-Based Criteria and Taxon-Specific Criteria. The EPA Committee envisions that Problem Formulation will be the critical process by which EPA will determine which type(s) of criteria will be needed for a given stressor (chemical) and that when it is not clear that one type is clearly the best, several types may be derived and compared. However, the EPA Committee has organized its subcommittees (topical working groups) around these three criteria types and each of these subcommittees is working on any/all of the aforementioned issues that are relevant to their criteria type. The following summaries provide the focus for each of the criteria types.

Water-based Criteria: The Water-based Criteria guidelines will focus on issues related to developing criteria: 1) for chemicals for which water concentration is a reasonable predictor of effects expected under natural exposure conditions, and 2) that will be derived and expressed as water concentrations.

Tissue-based Criteria: The Tissue-based Criteria guidelines will focus on issues related to developing criteria for chemicals for which water concentration is not a reasonable predictor of effects expected under natural exposure conditions (e.g. chemicals that bioaccumulate and/or biomagnify; chemicals for which diet is an important exposure pathway), and 2) that will be derived and expressed as either tissue concentrations and/or as water concentrations (by incorporating appropriate tissue-to-water translation procedures).

Taxon-Specific Criteria: The Taxon-specific Criteria guidelines will focus on techniques for modifying the aforementioned criteria types to provide appropriate levels of protection for specific taxa.

Charge Question 1.2. Please comment on whether the proposed criteria types and the scientific focus for each criteria type are logical and scientifically valid for developing a holistic and integrated criteria framework.

EPA is considering use of kinetic toxicity models for translating from (a) the constant exposure regimes of particular durations as are obtained from toxicity tests, to (b) continuously variable exposures of any duration as would occur in the real world. EPA is considering two kinetic toxicity models. Both use a structure of accumulation/deposition or damage/repair. The two models differ in how they explain why some individuals will live, while other apparently similar individuals will die at a particular exposure concentration. The stochastic process model (a.k.a., hazard function model) assumes a toxic stress operating on identical individuals, whereas the deterministic process model (a.k.a., frailty model) assumes a toxic stress operating on individuals having inherently different sensitivities.

Water Based Criteria

Charge Question 2.1. Please comment on whether the kinetic toxicity models being considered by EPA are scientifically appropriate for use in deriving water-based criteria.

EPA is considering the use of population models in order to combine the outcomes of differing types of effects: that is, effects on survival, growth, and reproduction. Such modeling also provides a way to account for the persistence of reduced population numbers after a toxic event is over: that is, recovery time. EPA is favoring the use of life-stage-structured models, so as not to discard information concerning sensitivities of different life stages, but has not settled upon a particular manner in which to apply these models, with respect to density dependence or independence.

Charge Question 2.2. Please comment on whether the population models being considered by EPA are scientifically appropriate for use in deriving water-based criteria.

The proposed approach evaluates multiple species. Therefore, a model for aggregating effects across species is necessary in order to derive a criterion. Although the concept that the tested species represent the potential range of sensitivities that might exist in the real world remains cogent, the subcommittee is considering some technical changes to replace the Species Sensitivity Distribution model used in the 1985 Guidelines.

Charge Question 2.3. Please comment on whether the proposal for aggregating effects across species being considered by EPA is scientifically appropriate for use in deriving water-based criteria.

EPA is considering some significant changes in the framework used to derive criteria. These changes are intended to improve the consideration of time-variable concentrations and better account for differences between life stages and between effects on survival, growth, and reproduction. The changes involve the use of more types of models during the derivation of criteria. The framework under consideration links a kinetic toxicity model to a population model to examine the impacts on each of several species being evaluated, and aggregates the effects across species into a measure of risk for one or more assemblages of species.

Charge Question 2.4 Please comment on whether the framework being considered by EPA for deriving water-based criteria is scientifically appropriate for use in deriving the criteria.

Tissue-Based Criteria

For chemicals with a high propensity to bioaccumulate in aquatic food webs and for which diet is a primary route of exposure, the EPA proposes to develop tissue-based criteria expressed as the chemical concentrations in specific animal tissues or dietary concentrations, with a process for translating to corresponding water and sediment concentrations. Tissue-based criteria allow for integration of multiple exposure pathways (water, diet) and facilitate direct comparison with environmental tissue concentrations to determine if there is a risk of adverse effects.

Charge Question 3.1. Please comment on the rationale and conceptual approach used for the development of tissue-based criteria for this group of chemicals. Is the SAB aware of other approaches for deriving criteria for these bioaccumulative chemicals that EPA should consider?

The proposed process for Tissue-based Criteria is intended to be flexible to maximize the use of available data and to accommodate certain limitations in the quality and quantity of data. This approach will also provide opportunities for states and tribes to develop alternative options that may be more suitable to site-specific conditions. National-level criteria may use deterministic approaches to characterize toxicity data when data are limited or probabilistic approaches (e.g., species sensitivity distributions) when data are sufficient. The process will also describe how a criterion may be refined on a site-specific basis when additional data are available.

Charge Question 3.2. Considering the strengths and limitations of the more flexible approach used to derive tissue-based criteria, please comment on the rationale and preference for allowing flexibility in the procedures used?

Unlike the dynamic exposure scenarios being addressed in development of water-based criteria, EPA is considering a steady-state approach for developing national criteria for bioaccumulative chemicals (i.e., modeling bioaccumulation and toxicity as a function of constant concentrations). Rationale for this approach is the much slower accumulation kinetics generally associated with these chemicals in higher trophic level fish and aquatic-dependent wildlife and concerns over their long-term bioaccumulation. In the context of population modeling, there appears to be much less residue-response information available for integrating responses of various demographic parameters over multiple life stages, such as fecundity and adult, juvenile, and larval survival. Consequently, it is not clear whether it would be feasible or useful to integrate population modeling into national-level tissue criteria for bioaccumulative chemicals. Current thinking is that where sufficient data exist to characterize exposure, bioaccumulation and toxicity on a dynamic basis, population modeling may evolve into an important tool in the development of site-specific criteria

Charge Question 3.3. Please comment on the rationale used by EPA for determining if/when to use population modeling in the development of Tissue-Based Criteria?

Taxon-Specific Criteria

EPA considers problem formulation an essential step of taxon-specific criteria development. Several considerations for problem formulation have been outlined in the proposed framework for deriving taxon-specific criteria under the categories of 1) appropriate taxonomic level for analysis, 2) means of data analysis, 3) different levels of protection, and 4) the taxon's or population's ability to tolerate adverse effects.

Charge Question 4.1. Please comment on the considerations for problem formulation outlined in the proposed framework for deriving Taxon-specific Criteria, specifically whether it will lead to scientifically defensible numeric criteria?

Toxicity data for special status species (e.g. federally-listed threatened or endangered; State-designated special status, etc.) are often quite limited, making consideration of other methods to predict effect concentrations for these taxa necessary. Several approaches utilizing surrogate species information have been identified: interspecies correlation estimates, species sensitivity distributions, identification of the most closely related species in the toxicological database for a specific chemical, and identification of the most sensitive species in the toxicological database for a specific chemical.

Charge Question 4.2. Of the approaches outlined for addressing surrogacy and gap analyses with regard to special status species, are there improvements to these tools

that would provide more scientifically defensible numeric criteria where specific data are not available? Are these tools adequate for developing scientifically defensible numeric criteria? What other tools are available to provide more scientifically defensible criteria when there is an absence of toxicological data for a specific pollutant and taxon?