COMMENTS SUMMARY REPORT

EXTERNAL PEER REVIEW OF
"EVALUATION OF THE CLEAN WATER ACT SECTION [CWA] 304(A) HUMAN HEALTH CRITERION FOR METHYLMERCURY: PROTECTIVENESS FOR THREATENED AND ENDANGERED [T&E] WILDLIFE IN CALIFORNIA"

Prepared for:

U.S. Environmental Protection Agency
Office of Water
Office of Science and Technology
Health and Ecological Criteria Division
1301 Constitution Ave., N.W.
Washington, D.C. 20004

Prepared by:

Versar, Inc.
6850 Versar Center
Springfield, Virginia 22151

September 2005

TABLE OF CONTENTS

I. INTRODUCTION

On May 18, 2000, EPA promulgated the California Toxics rule (CTR) for the State of California. The CTR included aquatic life and human health water quality criteria for priority toxic pollutants for California's inland surface waters, and enclosed bays and estuaries, because California had failed to timely adopt the criteria as required by the Clean Water Act (CWA). After a lengthy consultation under the Endangered Species Act (ESA), on March 24, 2000, the U.S. Fish and Wildlife Service (FWS) and the U.S. National marine Fisheries Service (collectively, the Services) issued a joint Biological Opinion which included an Incidental Take Statement and associated Reasonable and Prudent Measures for several pollutants. Terms and Conditions to implement the Reasonable and Prudent Measure for mercury included requiring EPA to revise its CWA 304(a) human health criteria guidance to ensure protection of Federally listed threatened and endangered aquatic and aquatic-dependant species (T&E species) in California, and to propose the revised criteria into the CTR.

EPA published its revised 304(a) human health mercury criteria guidance in January of 2001. The criterion is a fish tissue value of 0.3 mg of methylmercury per kilogram of fish tissue (wet weight). Since the Services Biological Opinion anticipated a revised water column value, an ESA consultation on the fish tissue criterion was initiated, and discussion started concerning whether the fish tissue criterion would be protective of California's T&E species.

In the fall of 2003, a team including mercury experts from the FWS and EPA's Office of Research and Development (ORD) was assembled to develop a methodology and perform analyses to determine if the revised 304(a) methylmercury criterion would be protective of California T&E species. The subject document reflects this group's methodology and analyses. Region 9 Water Division needs formal peer review of this document that was jointly crafted by EPA and FWS. The document establishes a methodology for determining whether fish tissue water quality criteria for bioaccumulative pollutants will adversely affect wildlife. As a pilot case, the analysis looked at the effects of EPA's new human health methylmercury water quality criterion on California's T&E species. The pilot case is intended to help satisfy the ESA requirement to develop a Biological Evaluation of the effects on California's T&E species from EPA's promulgation of its new methylmercury water quality criterion for California inland surface waters and enclosed bays and estuaries.

It is anticipated that the document will be the basis for justifying application of the water quality criterion to specific trophic level fish. In accordance with Agency policy, the document should be externally peer reviewed prior to moving forward with the promulgation of the methylmercury criterion for California waters.

Peer review is an important component of the scientific process. It provides a focused, objective evaluation of the document or material submitted for review. The criticism, suggestions and new ideas provided by the peer reviewers stimulate creative thought, strengthen the reviewed document and confer credibility on the product. Comprehensive, objective peer review leads to good science and product acceptance within the scientific community. This document was peer reviewed by Dr. Elwood Hill, Dr. Mark Sandheinrich, Dr. Edward Swain, and Dr. Marti Wolfe.

II. CHARGE TO THE PEER REVIEWERS

1. Issues/Questions/Concerns:

General Approach:
The methodology used for the risk assessment of non-fish species was based on exposure of methylmercury through ingestion of methylmercury contaminated prey. For fish species, risk was based on adverse effects associated with methylmercury tissue concentrations.

Question 1. Is this general approach reasonable? Why or why not?

Methodology - Bioaccumulation/Biomagnification:
The analysis assumes that although aquatic food webs vary considerably across all waters of California, the aquatic organisms can generally be classified into four trophic level categories. The analysis further assumes that biomagnification of methylmercury among trophic levels is sufficiently consistent across waterbodies that food chain multipliers can be estimated for this analysis based on draft national bioaccumulation factors (BAFs). If more site-specific BAF data exist, then that data can (and should) be used in place of the draft national BAFs to determine better site specific multipliers.

Question 2. Are the assumptions made concerning bioaccumulation and biomagnification between trophic levels reasonable for a statewide scale model and analysis? Why or why not?

Question 3. Do you believe another method (or a variation of this method) might present better results? Why or Why not? Stated another way: are there any elements of this analysis missing which you believe need to be included or would otherwise strengthen the analysis?

The State of California's Regional Water Quality Control Board used a similar analysis to determine numeric targets to protect wildlife in the Cache Creek and the Sacramento-San Joaquin Delta watersheds. Instead of determining fish tissue targets based on trophic levels, the Board developed targets based on specific size classes and species of fish that are believed to be consumed by the specific species of concern. That is, instead of deriving simple food chain multipliers based on the biomagnification expected when fish from one trophic level consume prey from the next lower trophic level, the Board derived trophic level ratios (TLRs) based on concentrations in fish prey that are similar in size (no-predator-prey interaction) but occupy different trophic levels. This discussion/explanation is found at pages 5 and 6 of the document entitled, "Evaluation of Numeric Wildlife Targets for Mercury in the Development of Total Maximum Daily Loads for the Cache Creek and Sacramento-San Joaquin Delta Watersheds".

Question 4. What are the reviewers comments concerning the use of TLRs to enhance the analysis when determining targets for wildlife species that feed from multiple trophic levels?

Methodology - Risk Assessment:
The risk assessment for non-fish (piscivorus wildlife) species included determining wildlife values (WVs) for particular species thought to be most at risk through dietary exposure. The WV represents the overall dietary concentration of methylmercury necessary to keep the daily ingested amount at or below a sufficiently protective reference dose, analogous to the tissue residue criterion for human health, and is

determined for selected species of concern based on body weight, total daily food ingestion rate and a protective reference dose:

$$WV = \frac{\text{Risk Reference Dose (RfD) X Body Weight of Species of Concern (BW)}}{\text{Summation of Food Ingestion Rate (FIR) (at each trophic level)}}$$

Reference Doses: Appropriate RfDs were determined for avian and mammalian species by first finding the most appropriate test dose, and then conducting uncertainty analyses.

Test Doses: The scientific literature was searched for all available oral test doses demonstrating observable effects concentrations, using the data preferences as outlined in the Great Lakes Initiative (GLI) Technical Support Document for Wildlife Criteria (page 9). Particular attention was paid to studies completed since the development of the GLI. The results and discussion of these searches are at pages 10 through 17. For a mammalian test dose, a NOEL dose of 0.33 mg/kg in food (0.055 mg/kg - bw/day) for the 145 day study by Wobeser et al. (1976a) was chosen as the appropriate test dose in this evaluation (page 12). For an avian test dose, a LOAEL of 0.064 mg/kg - bw/day calculated with Heinz control group feeding rate was chosen as an appropriate test dose in this evaluation (page 17).

Uncertainty Analyses: Interspecies Uncertainty Factors, Subchronic-to-Chronic Uncertainty Factors, and LOAEL-to-NOAEL Uncertainty Factors were considered for each of the mammalian and avian test doses. To convert the mammalian test dose and the avian test dose to reference doses, the GLI analysis and the Mercury Study Report to Congress (MSRC) (U.S. EPA 1997) analysis were considered and compared. This part of the analysis is at pages 17 -22.

Body Weights: The Body Weights for each species of concern were those of adult females, since the analysis assumed that the most sensitive endpoints relate to reproduction (reference doses were based on preventing adverse impacts to the reproductive viability of the species)**.** The studies used are included in the discussion at pages 24 - 46 for individual species.

Food Ingestion Rates: Dietary composition for specific species, and species-specific food ingestion rates (FIRs) are also discussed at pages 24 - 46. All available literature was reviewed for species-specific information. If sufficient species-specific information was not available, FIRs were determined using the most appropriate allometric equations from Nagy (2001).

Question 5. Are you aware of any other relevant and appropriate studies for test doses, body weights, food ingestion rates (including dietary composition) that should be considered and included in the discussion? What are they and why are they appropriate and relevant?

Question 6. Are the conclusions reached reasonable based on the information considered? Why or why not?

The calculated WVs for each species of concern were then compared with each species predicted dietary concentrations (DC). The DC represents an overall concentration of mercury in a species diet, resulting from application of the 0.3 mg/kg human health criterion. Two sets of DCs were determined, based on the trophic level (TL) mercury concentrations expected under each application approach (Average Concentration TL or Highest TL Approach) and the TL composition of the species' diet. WVs were then compared to predicted DCs. If a DC was higher than the WV, it was assumed that the species would likely have a dietary exposure under that application approach that may place it at risk for adverse effects from methylmercury toxicity.

Question 7. Is this approach reasonable? Why or Why not? Are there any elements of the analysis missing which you believe need to be included or would otherwise strengthen the analysis? If so, please explain.

In General:

Question 8. What is/are the weakest part(s) of the analysis and why? Do you have any suggestions for alternatives? Please discuss any other general or specific concerns you might have regarding any aspect of the analysis.

III. GENERAL COMMENTS

Mark Sandheinrich

General Comments and Overall Impression of the Scientific Merit of the Document

As a screening risk assessment, this study indicates that the proposed fish tissue criterion of 0.3 mg methylmercury/kg is likely to pose an ecological risk to some species of threatened and endangered wildlife in California. This suggests that further analysis and a more definitive ecological risk assessment is needed to determine specifically what hazard, if any, the proposed fish tissue criterion poses to threatened and endangered wildlife. The document does an excellent job critically reviewing the literature and using the best available information to derive model parameters. Moreover, several astute observations about methylmercury transfer between trophic levels and between adult birds, adult fish and their respective eggs indicates that the document authors are knowledgeable about factors influencing methylmercury bioaccumulation and toxicity.

The evaluation's shortcomings are due to the nature of the risk assessment methodology used (hazard quotient) and not necessarily through the application of the data to the process. Although appropriate for a screening-level risk assessment, the hazard quotient method does not provide a probabilistic estimate of risk and any policy decisions based solely on this assessment method may be subject to criticism because of the multiple assumptions that are required, including those with unknown scientific validation (i.e., uncertainty factors; Moore et al. 2003).

Several studies have taken a probabilistic approach to assessing the effects of methylmercury on wildlife (Duvall and Barron 2000, Hope 2003, Lohman et al. 2000, Moore et al. 1999, 2003) and this may be a more appropriate method than the method used in this document to evaluate the protectiveness of the methylmercury human health criterion for protecting endangered and threatened species of wildlife. The advantage of a probabilistic approach over the hazard quotient method is that uncertainty can be quantified, and the probability that exposure will exceed a NOAEL can be determined or, alternatively, probabilities of effects of varying magnitudes can be modeled.

Given the methodology chosen for the current evaluation, the document could be strengthened by (1) more explicitly stating the assumptions inherent in the model parameters, (2) using food chain multipliers derived specifically from California water bodies rather than from national mean BAFs, and (3) deriving multiple WVs with a wider range of uncertainty factors. If applicable, a probabilistic approach to evaluating the protectiveness for wildlife of the human criterion for methylmercury should be considered.

IV. RESPONSE TO CHARGE

General Approach:
The methodology used for the risk assessment of non-fish species was based on exposure of methylmercury through ingestion of methylmercury contaminated prey. For fish species, risk was based on adverse effects associated with methylmercury tissue concentrations.

1. Is this general approach reasonable? Why or why not?

Elwood Hill
The methodologies used in the document are **entirely reasonable**. Contrasted with human health where a single species is under constant study and review, wildlife health is difficult to evaluate if for no other reason than species and habitat diversity. Remarkably little toxicology in a "real world" sense is available for wildlife species. Due to the incredible complexity of natural systems, most field studies lack adequate controls and resources to provide truly relevant results beyond very limited objectives, and, of course, these studies rarely involve T & E species due to regulatory restraints. In conclusion, Dr. Russell et al. optimized their risk assessment in a superior fashion considering the dearth of relevant information in the peer-reviewed literature. Their projection of exposure through ingestion is the most usual source of MeHg for the species under review. However, alternative routes of exposure should not be over-looked in future risk assessments, e.g., inhalation, percutaneous.

Mark Sandheinrich
The risk assessment methodology for non-fish species was based solely on exposure of methylmercury through ingestion of methylmercury contaminated prey. This approach is reasonable and appropriate for two reasons:

1. Relative to uptake of methylmercury from aqueous sources, estimates of body normalized rates of water consumption for the species of concern and concentrations of methylmercury in surface freshwaters demonstrate that diet is the principal means of methylmercury uptake from aquatic systems for piscivorous birds and mammals.

2. Although several field studies have documented effects of methylmercury based on organ-specific or whole-organism concentrations, there was not an indication of the concentrations of methylmercury in the diet that resulted in those concentrations that caused the effect in the organism. Consequently, for this assessment, studies of methylmercury effects on non-fish species was limited to those in which the relationship between dose (i.e., dietary concentrations and consumption) and effect was determined.

Because the 304(a) human health mercury criterion is based on fish tissue concentrations, risk for fish species based on adverse effects associated with methylmercury tissue concentrations in most appropriate.

Edward Swain

Yes, this general approach is reasonable. The document acknowledges challenges and limitations, and deals with them in a thoughtful and robust manner. However, it was not an effortless application of the approach. The relevant toxicological studies were read carefully and in some cases the original authors were contacted for specific information that was not in the published results. Final judgment on the appropriate uncertainty factors for several of the endangered species required application of deep expertise on the part of the authors, and in the end, the decision was sometimes based on expert opinion. When this was necessary, the rationale for this expert opinion was well documented and presented.

Marti F. Wolfe

My overall assessment is that this approach is mostly reasonable, well researched and well documented, *given the current state of knowledge*. Therefore my comments are mostly address refinements, and/or additional specificity that can be used to strengthen the approach as results from studies currently being conducted become available.

Methodology - Bioaccumulation/Biomagnification:
The analysis assumes that although aquatic food webs vary considerably across all waters of California, the aquatic organisms can generally be classified into four trophic level categories. The analysis further assumes that biomagnification of methylmercury among trophic levels is sufficiently consistent across waterbodies that food chain multipliers can be estimated for this analysis based on draft national bioaccumulation factors (BAFs). If more site-specific BAF data exist, then that data can (and should) be used in place of the draft national BAFs to determine better site specific multipliers.

2. Are the assumptions made concerning bioaccumulation and biomagnification between trophic levels reasonable for a statewide scale model and analysis? Why or why not?

Elwood Hill
The assumptions made concerning bioaccumulation and biomagnification between trophic levels are reasonable in principle for a statewide scale model and analysis. However, and particularly in regard to T & E species local residue data (depending on the species range and habitat association) should be used in the risk assessment. This requirement may not always be possible when urgency affects response time, and, of course, many areas (even near significant point-source contamination) simply have not been sampled for residue analysis. Another issue is that most areas, even critical habitats, are not thoroughly sampled for residues and a common complaint is that the sampling design is weak. With this understanding the risk assessment must be based on the best information available. Thus, uncertainties must be invoked.

Mark Sandheinrich
Bioaccumulation Factors (BAFs) were used to derive food chain multipliers between trophic levels (TL) 2 and 3 and TL 3 and 4. These food chain multipliers were, in turn, used (1) to calculate TL concentrations necessary to maintain the overall human dietary concentration (DC) of 0.3 mg/kg in the Average Concentration Trophic Level Approach; (2) expected concentrations in TL 2 and TL 3 if the limiting methylmercury concentration is 0.3 mg/kg in trophic level 4 fish (i.e., Highest Trophic Level Approach) and (3) predicted dietary concentrations for each of the threatened and endangered wildlife species considered under the Average Concentration Trophic Level Approach and the Highest Trophic Level Approach.

The BAFs were from draft national BAFs presented in the EPA's methylmercury criterion document (Table 1) and presented here for discussion.

Table 1. Draft BAFs for methylmercury empirically derived from field data collected across the United States and reported in the open literature. Based on BAFs calculated from lotic and lentic systems.

	BAF Trophic Level 2	BAF Trophic Level 3	BAF Trophic Level 4
5th Percentile	18,000	74,000	250,000
Draft national values (approx geometric mean, 50th percentile)	120,000	680,000	2,700,000

95th Percentile	770,000	6,200,000	28,000,000

The document currently under review (page 5), used the approximate geometric mean BAF for each trophic level to calculate the food chain multipliers, based on (I assume) the assumption that the geometric mean best represents the wide range of BAFs from across the nation. From Table 1, and as the EPA acknowledges in the criterion document, it is evident that the range for BAFs for each trophic level varies by at least 10 (TL 2) to approximately 100 fold (TL 3 and TL 4). Moreover, the criterion document states "EPA fully recognizes that the approach taken to derive mercury BAFs collapses a very complicated non-linear process, which is affected by numerous physical, chemical, and biological factors, into a rather simplistic linear process. EPA also recognizes that uncertainty exists in applying a National BAF universally to all water bodies of the United States. Therefore, in the revised 2000 Human Health Methodology (EPA , 2000) we encourage and provide guidance for States, Territories, Authorized Tribes, and other stakeholders to derive site-specific field-measured BAFs when possible. In addition, should stakeholders believe some other type of model may better predict mercury bioaccumulation on a site-specific basis they are encouraged to use one, provided it is scientifically justifiable and clearly documented with sufficient data" (page A-18)

Using a method similar to that employed in the document under review to calculate food chain multipliers and the 5^{th} and 95^{th} percentiles for the BAFs, results in lower and upper bounds of the range of the food chain multiplier for TL 4 of at least

250,000/74,000 = 3.4

and 28,000,000/6,200,000 = 4.5

In turn, the lower and upper bounds of the range of the food chain multiplier for TL 3 is at least

74,000/18,000 = 4.1

and 6,200,000/770,000 = 8

Although, the range of the calculated food chain multipliers is not as great as that of the BAFs from which they are derived, based on the EPA's admission of the limitation of the draft national BAFs and the importance of the food chain multipliers to the risk assessment in question, the use of empirically derived national BAFs may not be appropriate for a statewide scale model and analysis, especially considering that BAFs from aquatic systems in California are available and could be used in the assessment.

Edward Swain

he assumptions are reasonable for assessing the potential for harm from methylmercury exposure, but not for assessing the exact effect on a given species. The document did not attempt to do the latter, so the modeling effort was appropriate. Even though there is a great deal of data represented in this document, and many calculations are taken to three decimal places, in actuality this is a coarse assessment of potential harm. I do not see evidence that the assessment procedure is consistently tilted toward conservative assumptions, which might predict harm when none would exist. Rather, I see that this approach is utilizing average conditions. For instance, food chain multipliers were calculated from the draft national BAFs presented in EPA's 2001 methylmercury criterion document (discussed in the text on page 5, but not listed in the references for some reason). These national BAFs are averages, but the 90% ranges in the BAFs have at least an order of magnitude of variation (p. A-14 of the 2001 document). The text on page 5 goes on to state that if site-specific BAF data are available, they could be used. I think that the point is that the use of average BAF data is reasonable unless good site-specific BAF data are available. But good site-specific BAF data are rarely available, and the potential for harm can be assessed using average data on a statewide scale.

Marti F. Wolfe

The analysis assumes that although aquatic food webs vary considerably across all waters of California, the aquatic organisms can generally be classified into four trophic level categories. This is reasonable as long as seasonal variation, migrational and/or opportunistic variations in trophic status are taken into consideration.

The analysis further assumes that biomagnification of methylmercury among trophic levels is sufficiently consistent across waterbodies that food chain multipliers can be estimated for this analysis based on draft national bioaccumulation factors (BAFs). If more site-specific BAF data exist, then that data can (and should) be used in place of the draft national BAFs to determine better site specific multipliers.

Draft national bioaccumulation factors are useful for calculations where site-specific data are lacking, but local derived factors are preferable. These can vary considerably even within a relatively homogeneous water system (Hope 2003).

Where more site-specific data do not exist, they should be sought, especially in habitats/systems where mercury behavior is highly variable and/or those impinging on threatened and endangered species, such as the Bay-Delta system.

Important work on trophic transfer on Bay and Delta marshlands has recently been completely by Letitia Grenier at UC Berkeley, and will soon be available in citable form.

Variation in California due to differences in Hg inputs, temperature, salinity and in methylation regimes, as mentioned above, can result in variability in biomagnification and bioaccumulation. As recommended by Weiner et al, habitats areas and trophic pathways associated with bioaccumulation and biomagnification of methylmercury to elevated concentrations should be identified (Weiner et al. 2003).
Furthermore, several mercury-bearing wetland areas of California are scheduled for restoration. This will provide much-need additional habitat for T & E species, but it is recognized that these activities may result in temporary elevations or 'pulses' of available and reactive mercury into food chains (Slotton et al. 2001). These areas will need to be monitored to determine if TRCs calculated pre-restoration are adequately protective during transient perturbations (Caldwell 2004).

3. Do you believe another method (or a variation of this method) might present better results? Why or Why not? Stated another way: are there any elements of this analysis missing which you believe need to be included or would otherwise strengthen the analysis?

Elwood Hill

The Russell method is superior to other approaches considering the type and quality of data drawn upon. With this understanding, the present approach to risk assessment is extremely thorough and should effectively allow sound decisions regarding potential MeHg hazard to the selected T & E species chosen for this evaluation. This document should serve as a model for future risk assessments fro both T & E and non-listed species of concern.

Mark Sandheinrich

The appropriateness of use of the draft national BAFs could be evaluated by (1) calculating BAFs for lakes and streams in California and (2) comparing the draft national values to the 95% CI (for example) for the values from California.

I believe that there are several variations of the method used in the risk assessment that might present "better" results. These are listed in order of predicted increasing precision and accuracy.

1. The document used the geometric mean national BAFs and then calculated the food chain multipliers for each trophic level. A better approach may be to calculate food chain multipliers for TL 3 and TL 4 for <u>individual</u> lakes and streams from the national data and <u>then</u> obtain the geometric mean food chain multipliers for each trophic level. The geometric mean values from the draft national BAFs for each TL probably are probably not derived from the same set of lakes and stream. This will undoubtably remove water bodies from the analysis, but will probably result in better estimates of food chain multipliers between trophic levels by using data from lakes and streams where BAFs for both trophic levels (used in calculating food chain muliplier) were actually derived.

2. Determine statewide geometric mean BAFs from data for lentic and lotic systems in California only. California lakes and streams are included in the national data set.

3. Same as (1), but use only lakes and streams from California.

Edward Swain

Better results would of course be produced if site-specific data were available—data associated with the species of interest. But, in the absence of site-specific data, the RWQCB's approach described in USFWS (2004) document is a desirable refinement. However, I feel that the distinctions made between the food chain multiplier (FCM) and trophic level ratio (TLR) is largely semantic, since the FCM is based on empirical trophic level data. The important point made in USFWS (2004) is the attention paid to the size of fish food items, since mercury concentration is very dependent on the size of a particular fish within a species. If there were any one change that could be made to the method that would be valuable, it would be adjusting the concentration of prey fish in a particular trophic level for the size of those fish. Some might argue that this is too fine a point, given the large variation already acknowledge within a trophic level. But, if a species is eating 100% of a small size fish in TL4, it could easily make a 3-fold difference in assumed concentration, equivalent to a jump in the uncertainty factor. 3-fold changes might not make much difference for some contaminants, especially completely anthropogenic chemicals such as PCBs, but for a chemical with significant natural levels that is perhaps only a factor of 3 lower than current ambient concentrations, we cannot afford to be cavalier with 3-fold differences. In fact, only one species in this analysis has a DC to WV ratio of appreciably more than three (that is, 300%)—the Yuma clapper rail (962%), and then only when the UFAis 3, rather than 1 (Table 5). The Light-footed clapper rail ratio is 408% with a UFA of 3, but in the grand scheme of things, 400% is pretty similar to 300%. It is noted on page 81 that specific research is needed to determine exactly how vulnerable the light-footed clapper rail is. The document is judicious in its call for research to address weaknesses in the analysis.

Marti F. Wolfe

The State of California's Regional Water Quality Control Board used a similar analysis to determine numeric targets to protect wildlife in the Cache Creek and the Sacramento-San Joaquin Delta watersheds. Instead of determining fish tissue targets based on trophic levels, the Board developed targets based on specific size classes and species of fish that are believed to be consumed by the specific species of concern. That is, instead of deriving simple food chain multipliers based on the biomagnification expected when fish from one trophic level consume prey from the next lower trophic level, the Board derived trophic level ratios (TLRs) based on concentrations in fish prey that are similar in size (no-predator-prey interaction) but occupy different trophic levels. This discussion/explanation is found at pages 5 and 6 of the document entitled, "Evaluation of Numeric Wildlife Targets for Mercury in the Development of Total Maximum Daily Loads for the Cache Creek and Sacramento-San Joaquin Delta Watersheds".

The Average Concentration Trophic level and Highest Trophic Level > comparison approach is more sensitive, if more demanding to apply, than the Regional Board's.

One particular strength of this work compared to other comparable efforts is that fish are regarded as *targets* of toxicity, and not only *vectors* of toxicity; and a summary of methylmercury toxicity to fish is included.

Work in progress will provide improved data on the impacts of MeHg to sturgeon (white and green), Sacramento splittail, and Delta smelt. These should be incorporated as they become available.

Invertebrates, both as targets and vectors, are under-represented. They should be included because of their abundance in the food web, and because the impact of mercury in invertebrates is much less well studied. For example we can use a rule of thumb for vertebrates that 90 – 99 % of total mercury in tissues is methylmercury, but the same blanket metric cannot be employed for inverts. This could be important in calculating the dose of animals that feed heavily on inverts.

The State of California's Regional Water Quality Control Board used a similar analysis to determine numeric targets to protect wildlife in the Cache Creek and the Sacramento-San Joaquin Delta watersheds. Instead of determining fish tissue targets based on trophic levels, the Board developed targets based on specific size classes and species of fish that are believed to be consumed by the specific species of concern. That is, instead of deriving simple food chain multipliers based on the biomagnification expected when fish from one trophic level consume prey from the next lower trophic level, the Board derived trophic level ratios (TLRs) based on concentrations in fish prey that are similar in size (no-predator-prey interaction) but occupy different trophic levels. This discussion/explanation is found at pages 5 and 6 of the document entitled, "Evaluation of Numeric Wildlife Targets for Mercury in the Development of Total Maximum Daily Loads for the Cache Creek and Sacramento-San Joaquin Delta Watersheds".

4. What are the reviewers comments concerning the use of TLRs to enhance the analysis when determining targets for wildlife species that feed from multiple trophic levels?

Elwood Hill

The "California Approach" adds a degree of specificity to the analysis, which is desirable. However, considering each level of specificity overlaid on a risk assessment usually results in reduced sample sizes and thoroughness of sampling; resultant limitations may increase rather than decrease uncertainty. When considering the best available science, one must use the best data available rather than adding to theoretical idealism. In other words, at least initially, take the cautious and proven approach. Every overlay of complexity increases uncertainty if the residue sampling is not expertly designed. Thus, it is possible to overstate the precision of analysis and develop erroneous conclusions. This will be a concern until more sophisticated models are developed. However, depending on the size and specificity of the of the range and habitat type for a species, the California approach may work very well, at least in theory.

Mark Sandheinrich

In evaluating the use of TLRs, it is instructive to first examine the concerns in using food chain multipliers raised by the document (USFWS 2004) evaluating methylmercury TMDLs for the Cache Creek and Sacramento-San Joaquin Delta watersheds.

The food chain multiplier was used in the document currently under review (USFWS 2003) to determine methylmercury concentrations in various trophic levels consumed by humans that, when combined, correspond to overall dietary concentrations of 0.3 mg/kg. In essence, food chain multipliers establish a linkage value for average methylmercury concentrations.between trophic levels. The food chain mulipliers were derived from ratios of geometric mean BAFs for trophic levels 4, 3, and 2 from a national data base. For the Average Concentration Trophic Level Approach, these linkage values were then used in combination with data from the human health methylmercury criterion document on human consumption of fish from these trophic levels to calculate the concentrations of methylmercury in each trophic level necessary to maintain the overall dietary concentration of 0.3 mg/kg. The same linkage values were also used in the Highest Trophic Level Approach to calculate concentrations of methylmercury in trophic levels 3 and 2 if 0.3 mg/kg is the limiting concentration of trophic level 4 fish. The food chain multiplier can also be calculated from existing tissue methylmercury concentrations in fish from the various trophic levels.

USFWS (2004) on page 5, indicates that "Both of these approaches to determining food chain multipliers assume there is a direct consumption link between trophic levels, with methylmercury concentrations in the higher trophic level fish resulting from ingesting the concentrations found in fish from the next lower trophic level". I respectfully disagree. Based on how it is calculated, I suggest that the term "food chain multiplier" is a misnomer. In fact, using the ratio of BAFs results in a value that is more appropriately call a "trophic level multiplier". The linkage variable (i.e., food chain multiplier) for TL 4, for example, is based on the geometric mean BAFs for water to trophic level 4 fish and for water to trophic level 3 calculated from a national database of BAFs. No predator-prey relationship in these BAF values is inferred. As indicated in my response to question 3, I imagine that not even the same identical set of water bodies were used to calculate the BAFs for each trophic level. (I have not examined the national database.) Similarly, using average tissue concentrations in fish from different trophic levels to calculate the linkage ratio also makes no stated assumptions about predator-prey relationships. All members of the same fish species are probably assigned to trophic levels based on the generic diet of adult members of the species as fish demonstrate a fairly strong ontogenetic shift in diet—even TL 4 fish are zooplanktivorous or invertivorous at some stage of life. The average mercury concentration for the TL is undoubtably based on the average concentration of all fish, regardless of size, for the species assigned the TL designation.

However, I do agree with USFWS (2004) that it is more appropriate to derive and apply a linkage value to size classes of fish that are actually consumed by the wildlife--especially considering that the food chain multipliers are calculated from national mean BAFs. In fact, I suggest that the TLR that USFWS (2004) recommends is really a refinement of the food chain multiplier developed by USFWS (2003). The data needed to generate the TLRs for application to a statewide assessment may be difficult to obtain and, given the role that other variables have in influencing the WV and DC (see response to question 6), TLRs may not provide an additional level of refinement desired for the risk assessment. When data is available, derivation of TLRs for site-specific criteria would be preferable and is much more defensible than linkage values based on BAFs from a national data set which, as demonstrated in my response to question 3, incorporate the wide variation in water bodies from across the nation.

Edward Swain

The use of TLRs is desirable if it is a means to more realistically model the size of prey that are actually being consumed. I have already discussed this issue in my response to Question 3.

Marti F. Wolfe

Methodology - Risk Assessment:

WV. The calculations employed in this work are consistent with current methods for arriving at Wildlife Values. The method has undergone some evolution since introduced in the GLI, and may continue to do so as improved data for its input values becomes available.

Test Doses:

Mammalian

Although the 30-year-old results of Wobesor were the best available when this document was written, newer studies on mink have recently been completed at the Nova Scotia Agricultural College using ecologically relevant concentrations of Hg (Nil Basu, pers. comm., McGill University, Montreal, Canada), and multiple sensitive endpoints (Basu et al. 2005, Basu 2005, Stamler et al. 2005). These data should be incorporated as they become available.

Avian

Work to refine and strengthen the avian WV is in progress, but less close to availability than the efforts for mink. Kenow, USGS, Meyer, Wisconsin Department of Environmental Quality and co-workers, will undertake an egg injection study in loons based on the methods of Heinz, in 2005 and 2006. Frederick and Spaulding, University of Florida, have currently under way a feeding study using captive great egrets. A WV using common loons has recently been presented by Evers and co-workers at Biodiversity Research Institute (Evers et al. 2003). Concurrent efforts of Schwarzbach and Suchanek will supply much-needed specifics from the Bay-Delta system; these will further enhance the precision of the model.

Uncertainty Analyses: The treatment of uncertainty analysis in this document reflects the most recent peer-reviewed sources for methylmercury, and is treated with thoroughness. However, recent work from other than wildlife species and substances other than mercury may be consulted to some benefit (Gray et al. 1998, Irwin et al. 2004, Lock and Smith 2003, Orphanides 2003, Snape et al. 2004, Waters and Fostel 2004).

Body Weights: For T & E species, the level of concern may need to shift from population-level to individual animal effects, in which case not only body weights, different metabolic responses of juveniles may need to be incorporated (Stark et al. 2004).

Food Ingestion Rates: Use of the Nagy reference represents the current standard. Further refinements may be obtained by consulted more recent work (McKechnie and Wolf 2004, White and Seymour 2004).

Methodology - Risk Assessment:
The risk assessment for non-fish (piscivorus wildlife) species included determining wildlife values (WVs) for particular species thought to be most at risk through dietary exposure. The WV represents the overall dietary concentration of methylmercury necessary to keep the daily ingested amount at or below a sufficiently protective reference dose, analogous to the tissue residue criterion for human health, and is determined for selected species of concern based on body weight, total daily food ingestion rate and a protective reference dose:

$$WV = \frac{\text{Risk Reference Dose (RfD) X Body Weight of Species of Concern (BW)}}{\text{Summation of Food Ingestion Rate (FIR) (at each trophic level)}}$$

<u>Reference Doses</u>: Appropriate RfDs were determined for avian and mammalian species by first finding the most appropriate test dose, and then conducting uncertainty analyses.

<u>Test Doses</u>: The scientific literature was searched for all available oral test doses demonstrating observable effects concentrations, using the data preferences as outlined in the Great Lakes Initiative (GLI) Technical Support Document for Wildlife Criteria (page 9). Particular attention was paid to studies completed since the development of the GLI. The results and discussion of these searches are at pages 10 through 17. For a mammalian test dose, a NOEL dose of 0.33 mg/kg in food (0.055 mg/kg - bw/day) for the 145 day study by Wobeser et al. (1976a) was chosen as the appropriate test dose in this evaluation (page 12). For an avian test dose, a LOAEL of 0.064 mg/kg - bw/day calculated with Heinz control group feeding rate was chosen as an appropriate test dose in this evaluation (page 17).

<u>Uncertainty Analyses</u>: Interspecies Uncertainty Factors, Subchronic-to-Chronic Uncertainty Factors, and LOAEL-to-NOAEL Uncertainty Factors were considered for each of the mammalian and avian test doses. To convert the mammalian test dose and the avian test dose to reference doses, the GLI analysis and the Mercury Study Report to Congress (MSRC) (U.S. EPA 1997) analysis were considered and compared. This part of the analysis is at pages 17 -22.

<u>Body Weights:</u> The Body Weights for each species of concern were those of adult females, since the analysis assumed that the most sensitive endpoints relate to reproduction (reference doses were based on preventing adverse impacts to the reproductive viability of the species). The studies used are included in the discussion at pages 24 - 46 for individual species.

<u>Food Ingestion Rates:</u> Dietary composition for specific species, and species-specific food ingestion rates (FIRs) are also discussed at pages 24 - 46. All available literature was reviewed for species-specific information. If sufficient species-specific information was not available, FIRs were determined using the most appropriate allometric equations from Nagy (2001).

5. Are you aware of any other relevant and appropriate studies for test doses, body weights, food ingestion rates (including dietary composition) that should be considered and included in the discussion? What are they and why are they appropriate and relevant?

Elwood Hill

I am not aware of other sources or authorities that would substantially improve the detail, care and logic of this risk assessment. Like most broad-based risk assessments, i.e., seven species, Russell et al. were subject to the quality and abundance of available data. They used the data very effectively in a well thought-out step-wise process in the evaluation of factors affecting selected T & E species in California. Their results and conclusions are defensible with the techniques serving as a sound model for future risk assessments. Further, their step-wise approach and attention to detail revealed weaknesses in data sets and will provide guidance for future data gathering. It is also appreciated that the authors respected and effectively built upon other risk assessment efforts, e.g., GLI, MSRC.

Mark Sandheinrich

I am unaware of other studies that should be considered to derive equation parameters for development of WV or DC values for the threatened and endangered species considered in this risk assessment. However, there are three recent studies by Jason Unrine and colleagues that are pertinent to the effects of methylmercury on amphibians.

In a laboratory study, Unrine and Jagoe (2004) fed larvae of southern leopard frogs (*Rana spenocephala*) aufwuchs with concentrations of inorganic and methylmercury that bracketed those observed in aufwuchs from the field and from sites reported in the literature to be contaminated with mercury from atmospheric deposition. They analyzed the metamorphs and tadpoles for inorganic and methylmercury. Inorganic Hg comprised the main source of Hg in the diets and the relative contribution of methylmercury to total mercury in the aufwuchs decreased with increasing concentration. The ratio of methylmercury to inorganic mercury in the tadpoles was greater than that of their diet and the BAF for methyl and inorganic mercury decreased with increasing concentration of mercury in the diet. A field study of mercury in frog gut contents and carcasses conducted in South Carolina wetlands (Unrine et al. 2005), confirmed their laboratory observations that amphibian larvae inhabiting areas contaminated with Hg by atmospheric deposition could be exposed to dietary mercury with concentrations as high as 1600 ng THg/ g dry weight. The majority of THg was present as inorganic Hg and the proportion of methylmercury decreased with increasing THg concentrations in the aufwuchs.

Based on their previous study of mercury bioaccumulation, Unrine et al. (2004) fed southern leopard frog larvae with aufwuchs that had mercury concentrations and speciation similar to those observed in aufwuchs from aquatic systems contaminated by atmospheric deposition. Dietary concentrations of 1500 – 3300 ng THg/g dry wt altered development and growth, and increased malformations and mortality of larvae. Methylmercury comprised 1.5-1.9% of the total mercury in the aufwuchs. They concluded that dietary mercury exposure in habitats contaminated by atmospheric deposition could adversely affect amphibian larvae.

There are few additional studies of the effects of methylmercury on fish relevant to this document. Houck and Cech (2004) found no effects of dietary methylmercury on bioenergetics of Sacramento blackfish (*Orthodon microlepidotus*) with approximately 2.5 mg/kg mercury in the muscle tissue. However, Drevnick and Sandheinrich (2003) found that dietary methylmercury at environmentally realistic concentrations suppressed estradiol and testosterone in female and male fathead minnows (*Pimephales promelas*), respectively. Suppressed levels of estradiol corresponded to altered gonad development in female fish and reduced spawning success. This study confirmed the results of a previous study by Hammerschmidt et al. (2002) that demonstrated that environmentally realistic levels of dietary methylmercury altered reproductive success of fathead minnows.

Edward Swain
No, I am not aware of any relevant studies that should have been considered for this analysis. The breadth of review and the thoughtful interpretation impressed me.

Marti F. Wolfe
I have included these under earlier questions.

6. Are the conclusions reached reasonable based on the information considered? Why or why not?

Elwood Hill

The conclusions are very well thought out and reasonable. As mentioned above, the risk assessment is orderly, detailed and well-presented. Again, per other risk assessments, the author was forced to use less than optimal data that were generated for a multitude of uses rather than having the luxury of a research plan for a defined purpose.

Mark Sandheinrich

In general, the conclusions reached are reasonable based on the risk assessment method employed and the data used to derive the RfD, FIR, and DC. If anything, the derived WVs may be underprotective because of the low UFs used to derive the RfD for each species. The range of the maximum cumulative UF (i.e., product of the individual UFs) used in determining the species RfDs was 3 -9. A review by Duke and Taggart (2000) of the UFs used in 24 screening ecological risk assessments in California reported that the maximum cumulative UF ranged from 10 to 3000. However, they noted a "paucity of agency guidelines for selecting UF for HQ *(hazard quotient*—my note*)* calculation and little information intended for agency use on appropriate UFs for varying conditions and extrapolations". In the current risk assessment, guidance on determining the appropriate values for each uncertainty factor was obtained from USEPA (1995) and Abt Associates (1995—same as Abt Associates reference below?) but these documents also give relatively minimal guidance for value derivation. For example, with regards to the interspecies uncertainty factor (UF_A), USEPA (1995) indicates that information on species sensitivity differences "provide support for the 1-100 recommended range for the value" but stress that "the actual selection of an interspecies UF for application to a particular situation must be made on a case-by-case basis and requires the use of *best professional judgement"* (page 20; emphasis mine). Abt Associates (1995) suggested that an interspecies uncertainty factor of 10 was justified (as cited in Chapman et al. 1998—I did not have access to Abt Associates 1995). Calabrese and Baldwin (1993) recommended an UF_A of at least 10 to protect endangered species.

As a further illustration, if an interspecies uncertainty factor of 10 is used, then the WV for the sea otter becomes 0.0055 mg/kg diet and the calculated DC for the Average Concentration Trophic Level Approach exceeds the WV by 10 fold and by more than 4 fold for the Highest Trophic Level Approach. Under this scenario, the EPA's human health TRC (0.3 mg/kg) is likely to result in a dietary exposure that would place sea otters at risk, contrary to the conclusion reached by the document under review (page 64). A similar analysis could be conducted for the other species. The document states that the interspecies uncertainty factor of 1 was selected because the southern sea otter is in the same taxonomic family (although different genera) as the mink from which the TD was determined and, therefore, should have similar sensitivity. However, if protection of individuals is the goal, then a careful assessment of intra- and interspecies extrapolations is warranted. For example, in the review by Chapman et al. (1998), they noted that the U.S. EPA commissioned a report by Abt Associates (1995) to support development of wildlife criteria for the Great Lakes water quality initiative. A review and analysis by Abt Associates (1995) of subchronic and chronic data from 174 toxicity studies with birds and mammals showed that 94% of the species sensitivities were within a factor of 50 of each other and *ranged 1 to 1000* (emphasis mine). Moreover, Chapman et al (1998) reported that, based on intraspecific variation in response by laboratory rodents in toxicity tests, human variability in toxicological assessments is assigned an uncertainty factor of 10 and "many ecotoxicologists assume that wildlife species tested for toxicity responses would have a greater intraspecific variation than inbred laboratory rodent strains because of greater genetic diversity". Hence, an interspecies uncertainty factor of 10 is not unreasonable.

My point here is not necessarily to criticize the values selected for each UF, nor advocate the use of a particular value. Rather, my intent is to highlight the fact, as others have more eloquently done, that the use of uncertainty factors (aka safety factors) makes some assumptions with unknown scientific justification and, hence, is perhaps one of the weakest links in an ecological risk assessment that uses hazard quotients.

For the current document under review, I concur with several of the recommendations made by Chapman et al. (1998) for the use of safety factors. Specifically, "(1) Extrapolation requires context. Any use of safety factors should be based on existing scientific knowledge and should include appropriate caveats; (2) Extrapolation is not fact. Safety factors should be used only for screening, not as threshold or absolute values; (3) Extrapolation is uncertain. Safety factors should encompass a range, not a single number". In addition, it may be appropriate to provide some indication of how the UFs used in this document compare to those used by the state for ecological risk assessment.

Edward Swain

The conclusions are reasonable, based on the information considered. The information is limited in many ways, and so it is appropriate to not over-interpret it by creating a misleading degree of detail. This analysis certainly reveals which species are more at risk, and therefore which species for which more data should be obtained.

I found the premise of the document a bit inefficient, from a scientist's point of view. I had to keep remembering that the endangered species were not being evaluated against present ambient conditions, but, rather, what if the mercury concentrations of fish were in compliance with the water quality criterion of 0.3 mg/kg of edible portions. And, then, to further complicate the scientific analysis, there were two different approaches within which this criterion might be met. I felt that this series of premises made the document difficult to read and unnecessarily difficult to determine if the conclusions were reasonable. However, after multiple readings, I am confident that I have followed the logic and find it reasonable. In fact, having been a peer reviewer of many scientific documents, I find this document uncommonly well written. I feel that the authors were given tight constraints within which to work, and that an admirable job was accomplished. A much more satisfying product would have been produced had site-specific information been made available to the authors, but such information was not available.

Marti F. Wolfe

Yes, the conclusions reached are reasonable, and the document as it stands is suitable for regulatory decision making. However, as results from studies currently underway become available, they should be incorporated - that is, the document should be considered as an evolving method, not a fixed and unchangeable entity.

The calculated WVs for each species of concern were then compared with each species predicted dietary concentrations (DC). The DC represents an overall concentration of mercury in a species diet, resulting from application of the 0.3 mg/kg human health criterion. Two sets of DCs were determined, based on the trophic level (TL) mercury concentrations expected under each application approach (Average Concentration TL or Highest TL Approach) and the TL composition of the species' diet. WVs were then compared to predicted DCs. If a DC was higher than the WV, it was assumed that the species would likely have a dietary exposure under that application approach that may place it at risk for adverse effects from methylmercury toxicity.

7. Is this approach reasonable? Why or Why not? Are there any elements of the analysis missing which you believe need to be included or would otherwise strengthen the analysis? If so, please explain.

Elwood Hill

The development and use of wildlife values for each species is critical for this risk assessment. The specificity is especially useful where individual species are of interest, i.e., T & E. The approach presented herein is clearly the best I am aware of for the intended purpose. Fortunately for this "bench-mark" risk assessment (though not for wildlife !), MeHg is a nearly universal contaminant that has received abundant attention. Therefore, the MeHg database is probably as complete as can be found for any contaminant. However, this model will have to be reconsidered for risk assessments involving labile contaminants or when the purpose of the risk assessment addresses diverse species in a critical habitat, even when the habitat of concern is well defined.

Regarding the use of the 0.3 mg/kg, ww human health criteria for wildlife, i.e., average concentration, I do not believe it is appropriate for wildlife, even though it appeared to be adequate for some T & E species. My concerns are twofold. 1) wildlife generally ingest the entire prey, not just the "edible portions" (though shellfish are a bit more relevant). To my knowledge there is no satisfactory correction from filet to whole carcass. 2) Philosophically, I do not want to lose the sophistication of the present MeHg risk assessment to a "canned" non-specific human health model that does not address the uniqueness of species diversity. Wildlife are exposed differently than humans despite the commonality of peroral ingestion. The average concentration model would successfully protect (theoretically) fewer than half of the species evaluated and would have left most of the others seriously "at risk." While regulatory agencies would prefer to have the perfect single variable risk assessment for humans and wildlife, it just cannot be. No matter how theoretically desirable, no single species can adequately represent even their close relatives (cf. California and Yuma clapper rails).

Mark Sandheinrich

he dual approach of using DCs based on the Average Concentration TL Approach and Highest TL Approach is reasonable and allows comparison of potential risk under the method used to derive the human health criterion (Average Concentration TL) and under the method by which most states would, I assume, apply the human health criterion (Highest TL Approach).

Edward Swain

Some might suggest that the analysis would be strengthened through the use of probabilistic risk assessment methods. This suggestion is attractive, in that it would yield quantitative estimates of the proportion of individuals in the population that are at what degree of risk. Such estimates would, however, give a false sense of knowledge, given the nature of the inputs to this analysis. A probabilistic assessment would be more appropriate with site-specific information, in contrast to the largely national average data used in this analysis. I am not completely happy with the exact numbers that are produced with the present approach, but the discussion that accompanies the exact numbers utilizes them appropriately—they are interpreted more qualitatively than quantitatively. Because this document presented two different approaches, the "Average Concentration Trophic Level Approach" and the "Highest Trophic Level Approach", and always discussed both results, there was a degree of temperance to the conclusions, such as in IX.B. (pages 81 to 82).

Marti F. Wolfe

There is insufficient attention given to data coming from outside the conventionally accepted feeding-study-in-a-wildlife species. Once again, this is less a reflection on the authors of this documents than a failure of vision of all of us in the field. Toxicogenomics, QSAR (quantitative structure activity relationships) inter-species toxicokinetics extrapolations, etc in the human health and laboratory animal arena have moved far beyond us in wildlife. Our failure to close this gap only deprives us of powerful tools which we can most efficaciously employ in risk assessment for T & E species but will benefit wildlife species in general.

In general:

8. What is/are the weakest part(s) of the analysis and why? Do you have any suggestions for alternatives? Please discuss any other general or specific concerns you might have regarding any aspect of the analysis.

Elwood Hill

The evaluation was thorough, logical and very well presented. Considering that the evaluation and comparisons were based on existing rather than designed data sets, the cooperators on this evaluation developed a defensible "bench-mark" risk assessment for its intended purpose of protecting individual T & E species. The only weakness in this study was the derivation of uncertainty values, and that certainly was not the fault of the collaborators.

Mark Sandheinrich

I think the analysis has shortcomings in several areas.

(1) Values for different model parameters are presented without careful statement and consideration of the assumptions that were made in deriving those parameters. This leaves the uninformed reader with the impression that the parameter values are deterministic and without variability (e.g., food chain multipliers).

(2) The UFs have a major role in determining whether or not the DCs will exceed the WVs and result in a conclusion of probable harm. The UFs values selected need stronger scientific justification other than that they were the same as those used in the Great Lake initiative and Mercury Study Report to Congress. The degree of conservatism of the risk assessment is unknown.

(3) The conclusions based on the risk assessment are deterministic without measure of the uncertainty or variability associated with model parameters or estimates of the probability of adverse effects to the species of concern. A probabilistic risk assessment model would better incorporate the uncertainty and stochasticity of model parameters and provide decision makers with better information on the probability of adverse effects to the threatened and endangered species.

Edward Swain

The weakest parts of the analysis are largely issues that the authors cannot control: Use of average BAFs because site-specific data was not collected; Unknown degree of methylmercury sensitivity of an untested species; overly broad dose spacing of the toxicity tests that were performed (NOAEL-LOAEL issues); Uncertainty in converting subchronic data to chronic. These weaknesses were acknowledged and adequately dealt with in the document, but certainly are still weaknesses. The USFWS (2004) document discusses a weakness that could conceivably be adequately dealt with even without a great deal of new data: within a trophic level, fish accumulate mercury in a predictable manner as they grow. To incorporate this knowledge in modeling the mercury concentration of food would increase accuracy of the modeling effort. If predators are restricted to a given size of fish within a trophic level, it would increase the accuracy of mercury exposure estimates to include that information in some quantitative manner.

Marti F. Wolfe

The weakest part of the analysis is that its most core component, the WV, is based on data ~30 years old. This is no reflection on the quality of the effort; it is simply the state of the science, one that everyone working in the field laments. The work in progress that I have described and/or referenced above will do much to add refinements and enhance predictability of MeHg risk assessment to wildlife species.

The in-progress mammal data promise to be more robust and more immediately applicable by virtue of coming from a single lab, as a coordinated effort.

This was also the case with the Heinz-Hoffman mallard studies of the 1970-s and is one reason those data have stood the test of time and continue to be useful for current risk assessments.

The avian work currently in progress, in contrast, represents the effort of a number of investigators in government, university, and NGO organizations. Although the quality of the data is anticipated to be high, there is reason for concern that it may not be a further refinement over Heinz-Hoffman because of differences in protocol, endpoints measured, etc.

I propose that a coordinating council for avian MeHg be established, that asks the following questions:

1) What data do we need besides Heinz-Hoffman for an improved avian MeHg WV?
2) Which of those data can be anticipated from studies currently under way?
3) What additional studies (species, endpoints) are needed to fill the gaps identified by the answer to #2?

Until those data are obtained and validated via the peer-review process, this document represents the best that we have. It is a very commendable work.

V. SPECIFIC COMMENTS

None.

VI. ADDITIONAL REFERENCES

Mark Sandheinrich

General

Duvall, S. E., and M. G. Barron. 2000. A screening level probabilistic risk assessment of mercury in Florida Everglades food webs. Ecotoxicology and Environmental Safety 47:298-305.

Hope, B. 2003. A basin-specific aquatic food web biomagnification model for estimation of mercury target levels. Environmental Toxicology and Chemistry 22:2525-2537.

Lohman, K., P. Pai, C. Seigneur, D. Mitchell, K. Heim, K. Wandland, and L. Levin. 2000. A probabilistic analysis of regional mercury impacts on wildlife. Human and Ecological Risk Assessment 6:103-130.

Moore, D. R. J., B. E. Sample, G. W. Suter, B. R. Parkhurst, and R. S. Teed. 1999. A probabilistic assessment of the effects of methylmercury and PCBs on mink and kingfishers along East Fork Poplar Creek, Oak Ridge, Tennessee, USA. Environmental Toxicology and Chemistry 18:2941-2953.

Moore, D. R. J., R. S. Teed, and G. M. Richardson. 2003. Derivation of an ambient water quality criterion for mercury: taking account of site-specific conditions. Environmental Toxicology and Chemistry 22:3069-3080.

Question 2

U. S. EPA. 2001. Water quality criterion for the protection of human health: methylmercury. EPA-823-R-01-001. Office of Science and Technology, Office of Water, U.S. Environmental Protection Agency Washington, DC.

Question 5

Drevnick, P.E. and Sandheinrich, M.B. 2003. Effects of dietary methylmercury on reproductive endocrinology of fathead minnows. Environ. Sci. Technol. 37: 4390-4396.

Houck,A., and J. J. Jr. Cech. 2004. Effects of dietary methylmercury on juvenile Sacramento blackfish bioenergetics. Aquatic Toxicology (69) 107-123

Unrine, J.M., and C. H. Jagoe. 2004. Dietary mercury exposure and bioaccumulation in southern leopard frog (*Rana sphenocephala*) larvae. Environmental Toxicology and Chemistry 23: 2956–2963.

Unrine, J.M., C. H. Jagoe, W. A. Hopkins, and H. A. Brant. 2004. Adverse effects of ecologically relevant dietary mercury exposure in southern leopard frog (*Rana sphenocephala*) larvae. Environmental Toxicology and Chemistry 23: 2964–2970.

Unrine, J. M., C.H. Jagoe, A.C. Brinton, H.A. Brant, and N.T. Garvin. 2005. Dietary mercury exposure and bioaccumulation in amphibian larvae inhabiting Carolina bay wetlands. Environmental Pollution. In Press.

<u>Question 6</u>

Abt Associates. 1995. Technical basis for recommended ranges of uncertainty factors used in deriving wildlife criteria for the Great Lakes water quality initiative. Final Report. Office of Water, U. S. Environmental Protection Agency, Washington, DC.

Calabrase, E. J., and L. A. Baldwin. 1993. Performing ecological risk assessments. Lewis, Chelsea, MI, USA.

Chapman, P. M., A. Fairbrother, and D. Brown. 1998. A critical evaluation of safety (uncertainty) factors for ecological risk assessment. Environmental Toxicology and Chemistry 17:99-108.

Duke, L. D., and M. Taggart. 2000. Uncertainty factors in screening ecological risk assessments. Environmental Toxicology and Chemistry 19:1668-1680.

USEPA. 1995. Great Lakes Water Quality Initiative Technical Support Document for Wildlife Criteria. EPA-820-B-95-009. Office of Water. Washington, DC

Edward Swain

USEPA. 2001. Water Quality Criterion for the Protection of Human Health: Methylmercury. EPA-823-R-01-001. January 2001

USFWS. 2004. Evaluation of Numeric Wildlife Targets for Methylmercury in the Development of Total Maximum Daily Loads for the Cache Creek and Sacramento-San Joaquin Delta Watersheds. 30 pp.

Marti F. Wolfe

Basu, N., Stamler, C.J., Loua, K.M., and Chan, H.M. 2005. 2005. An inter-species comparison of mercury inhibition on muscarinic acetylcholine receptor binding in the cerebral cortex and cerebellum. Toxicology and Applied Pharmacology In press: doi:10.1016/j.taap.2004.09.009:

Basu, N., K. Klenavic, M. Gamberg, M. O'Brien, R. D. Evans, A. M. Scheuhammer and H. M. Chan. 2005. Effects of mercury on neurochemical receptor binding characteristics in wild mink. Environmental Toxicology and Chemistry 24: 39-45.

Caldwell, B. a. 2004. Mercury Technical Memorandum, South Bay SaltPond Restoration Project. San Francisco, CA. California State Coastal Conservancy U.S. Fish and Wildlife Service, California Department of Fish and Game.

Drevnick, P.E. and Sandheinrich, M.B. 2003. Effects of dietary methylmercury on reproductive endocrinology of fathead minnows. Environ. Sci. Technol. 37: 4390-4396.

Evers, D. C., O. P. Lane and L. Savoy. 2003. Assessing the impacts of methylmercury on piscivorous wildlife using a wildlife criterion value based on the Common Loon, 1998-2002. BioDiversity Research Institute, Falmouth. Maine.

Gray, L. E. J., J. Ostby, C. Wolf, C. Lambright and W. Kelce. 1998. The value of mechanistic studies in laboratory animals for the prediction of reproductive effects in wildlife: Endocrine effects on mammalian sexual differentiation. Environmental Toxicology And Chemistry 17: 109-118.

Hope, B. 2003. A basin-specific aquatic food web biomagnification model for estimation of mercury target levels. Environ Toxicol Chem 22: 2525-37.

Houck,A., and J. J. Jr. Cech. 2004. Effects of dietary methylmercury on juvenile Sacramento blackfish bioenergetics. Aquatic Toxicology (69) 107-123

Irwin, R. D., G. A. Boorman, M. L. Cunningham, A. N. Heinloth, D. E. Malarkey and R. S. Paules. 2004. Application of toxicogenomics to toxicology: basic concepts in the analysis of microarray data. Toxicol Pathol 32 Suppl 1: 72-83.

Lock, E. A. and L. L. Smith. 2003. The role of mode of action studies in extrapolating to human risks in toxicology. Toxicol Lett 140-141: 317-22.

McKechnie, A. E. and B. O. Wolf. 2004. The allometry of avian basal metabolic rate: good predictions need good data. Physiol Biochem Zool 77: 502-21.

Orphanides, G. 2003. Toxicogenomics: challenges and opportunities. Toxicol Lett 140-141: 145-8.

Slotton, D. G., S. M. Ayers, T. Suchanek, R. D. Weyand, A. M. Liston, C. Asher and D. C. Nelson. 2001. The Effects of Wetlands Restoration on the Production and Bioaccumulation of Methyl Mercury in The Sacramento-San Joaquin Delta, California. CALFED Proposal

Snape, J. R., S. J. Maund, D. B. Pickford and T. H. Hutchinson. 2004. Ecotoxicogenomics: the challenge of integrating genomics into aquatic and terrestrial ecotoxicology. Aquat Toxicol 67: 143-54.

Stamler, C. J., N. Basu and H. M. Chan. 2005. Biochemical markers of neurotoxicity in wildlife and human populations: Considerations for method development. Journal of Toxicology and Environmental Health In press; accepted: Dec 17, 2004:

Stark, J. D., J. E. Banks and R. Vargas. 2004. How risky is risk assessment: the role that life history strategies play in susceptibility of species to stress. Proc Natl Acad Sci U S A 101: 732-6.
Waters, M. D. and J. M. Fostel. 2004. Toxicogenomics and systems toxicology: aims and prospects. Nat Rev Genet 5: 936-48.

Unrine, J. M., C.H. Jagoe, A.C. Brinton, H.A. Brant, and N.T. Garvin. 2005. Dietary mercury exposure and bioaccumulation in amphibian larvae inhabiting Carolina bay wetlands. Environmental Pollution. In Press.

Unrine, J.M., C. H. Jagoe, W. A. Hopkins, and H. A. Brant. 2004. Adverse effects of ecologically relevant dietary mercury exposure in southern leopard frog (*Rana sphenocephala*) larvae. Environmental Toxicology and Chemistry 23: 2964–2970.

Unrine, J.M., and C. H. Jagoe. 2004. Dietary mercury exposure and bioaccumulation in southern leopard frog (*Rana sphenocephala*) larvae. Environmental Toxicology and Chemistry 23: 2956–2963.

Weiner, J. G., C. G. Gilmour and D. P. Krabbenhoft. 2003. Mercury Strategy for the Bay-Delta Ecosystem: A Unifying Framework for Science, Adaptive Management, and Ecological Restoration. California Bay Delta Authority.

White, C. R. and R. S. Seymour. 2004. Does basal metabolic rate contain a useful signal? Mammalian BMR allometry and correlations with a selection of physiological, ecological, and life-history variables. Physiol Biochem Zool 77: 929-41.

APPENDIX A

REVIEWER COMMENTS

PEER REVIEWER COMMENTS

DR. ELWOOD HILL

Peer review: E. F. Hill

General Approach:

#1. The methodologies used in the document are entirely reasonable. Contrasted with human health where a single species is under constant study and review, wildlife health is difficult to evaluate if for no other reason than species and habitat diversity. Remarkably little toxicology in a "real world" sense is available for wildlife species. Due to the incredible complexity of natural systems, most field studies lack adequate controls and resources to provide truly relevant results beyond very limited objectives, and, of course, these studies rarely involve T & E species due to regulatory restraints. In conclusion, Dr. Russell et al. optimized their risk assessment in a superior fashion considering the dearth of relevant information in the peer-reviewed literature. Their projection of exposure through ingestion is the most usual source of MeHg for the species under review. However, alternative routes of exposure should not be over-looked in future risk assessments, e.g., inhalation, percutaneous.

Methodology-Bioaccumulation/Biomagnification:

#2. The assumptions made concerning bioaccumulation and biomagnification between trophic levels are reasonable in principle for a statewide scale model and analysis. However, and particularly in regard to T & E species local residue data (depending on the species range and habitat association) should be used in the risk assessment. This requirement may not always be possible when urgency affects response time, and, of course, many areas (even near significant point-source contamination) simply have not been sampled for residue analysis. Another issue is that most areas, even critical habitats, are not thoroughly sampled for residues and a common complaint is that the sampling design is weak. With this understanding the risk assessment must be based on the best information available. Thus, uncertainties must be invoked.

#3. The Russell method is superior to other approaches considering the type and quality of data drawn upon. With this understanding, the present approach to risk assessment is extremely thorough and should effectively allow sound decisions regarding potential MeHg hazard to the selected T & E species chosen for this evaluation. This document should serve as a model for future risk assessments for both T & E and non-listed species of concern.

#4. The "California Approach" adds a degree of specificity to the analysis, which is desirable. However, considering each level of specificity overlaid on a risk assessment usually results in reduced sample sizes and thoroughness of sampling; resultant limitations may increase rather than decrease uncertainty. When considering the best available science, one must use the best data available rather than adding to theoretical idealism. In other words, at least initially, take the cautious and proven approach. Every overlay of complexity increases uncertainty if the residue sampling is not expertly designed. Thus, it is possible to overstate the precision of analysis and develop erroneous conclusions. This will be a concern until more sophisticated models are developed. However, depending on the size and specificity of the of the range and habitat type for a species, the California approach may work very well, at least in theory.

Methodology-Risk Assessment:

#5. I am not aware of other sources or authorities that would substantially improve the detail, care and logic of this risk assessment. Like most broad-based risk assessments, i.e., seven species, Russell et al. were subject to the quality and abundance of available data. They used the data very effectively in a well thought-out step-wise process in the evaluation of factors affecting selected T & E species in California. Their results and conclusions are defensible with the techniques serving as a sound model for future risk assessments. Further, their step-wise approach and attention to detail revealed weaknesses in data sets and will provide guidance for future data gathering. It is also appreciated that the authors respected and effectively built upon other risk assessment efforts, e.g., GLI, MSRC.

#6. The conclusions are very well thought out and reasonable. As mentioned above, the risk assessment is orderly, detailed and well-presented. Again, per other risk assessments, the author was forced to use less than optimal data that were generated for a multitude of uses rather than having the luxury of a research plan for a defined purpose.

#7. The development and use of wildlife values for each species is critical for this risk assessment. The specificity is especially useful where individual species are of interest, i.e., T & E. The approach presented herein is clearly the best I am aware of for the intended purpose. Fortunately for this "bench-mark" risk assessment (though not for wildlife !), MeHg is a nearly universal contaminant that has received abundant attention. Therefore, the MeHg database is probably as complete as can be found for any contaminant. However, this model will have to be reconsidered for risk assessments involving labile contaminants or when the purpose of the risk assessment addresses diverse species in a critical habitat, even when the habitat of concern is well defined.

Regarding the use of the 0.3 mg/kg, ww human health criteria for wildlife, i.e., average concentration, I do not believe it is appropriate for wildlife, even though it appeared to be adequate for some T & E species. My concerns are twofold. 1) wildlife generally ingest the entire prey, not just the "edible portions" (though shellfish are a bit more relevant). To my knowledge there is no satisfactory correction from filet to whole carcass. 2) Philosophically, I do not want to lose the sophistication of the present MeHg risk assessment to a "canned" non-specific human health model that does not address the uniqueness of species diversity. Wildlife are exposed differently than humans despite the commonality of peroral ingestion. The average concentration model would successfully protect (theoretically) fewer than half of the species evaluated and would have left most of the others seriously "at risk." While regulatory agencies would prefer to have the perfect single variable risk assessment for humans and wildlife, it just cannot be. No matter how theoretically desirable, no single species can adequately represent even their close relatives (cf. California and Yuma clapper rails).

In General:

#8. The evaluation was thorough, logical and very well presented. Considering that the evaluation and comparisons were based on existing rather than designed data sets, the cooperators on this evaluation developed a defensible "bench-mark" risk assessment for its intended purpose of protecting individual T & E species. The only weakness in this study was the derivation of uncertainty values, and that certainly was not the fault of the collaborators.

PEER REVIEW COMMENTS:

DR. MARK SANDHEINRICH

Evaluation of the Clean Water Action Section 304(a) human health criterion for methylmercury: protectiveness for threatened and endangered wildlife in California

Review by
Mark Sandheinrich
University of Wisconsin-La Crosse

General Comments and Overall Impression of the Scientific Merit of the Document

As a screening risk assessment, this study indicates that the proposed fish tissue criterion of 0.3 mg methylmercury/kg is likely to pose an ecological risk to some species of threatened and endangered wildlife in California. This suggests that further analysis and a more definitive ecological risk assessment is needed to determine specifically what hazard, if any, the proposed fish tissue criterion poses to threatened and endangered wildlife. The document does an excellent job critically reviewing the literature and using the best available information to derive model parameters. Moreover, several astute observations about methylmercury transfer between trophic levels and between adult birds, adult fish and their respective eggs indicates that the document authors are knowledgeable about factors influencing methylmercury bioaccumulation and toxicity.

The evaluation's shortcomings are due to the nature of the risk assessment methodology used (hazard quotient) and not necessarily through the application of the data to the process. Although appropriate for a screening-level risk assessment, the hazard quotient method does not provide a probabilistic estimate of risk and any policy decisions based solely on this assessment method may be subject to criticism because of the multiple assumptions that are required, including those with unknown scientific validation (i.e., uncertainty factors; Moore et al. 2003).

Several studies have taken a probabilistic approach to assessing the effects of methylmercury on wildlife (Duvall and Barron 2000, Hope 2003, Lohman et al. 2000, Moore et al. 1999, 2003) and this may be a more appropriate method than the method used in this document to evaluate the protectiveness of the methylmercury human health criterion for protecting endangered and threatened species of wildlife. The advantage of a probabilistic approach over the hazard quotient method is that uncertainty can be quantified, and the probability that exposure will exceed a NOAEL can be determined or, alternatively, probabilities of effects of varying magnitudes can be modeled.

Given the methodology chosen for the current evaluation, the document could be strengthened by (1) more explicitly stating the assumptions inherent in the model parameters, (2) using food chain multipliers derived specifically from California water bodies rather than from national mean BAFs, and (3) deriving multiple WVs with a wider range of uncertainty factors. If applicable, a probabilistic approach to evaluating the protectiveness for wildlife of the human criterion for methylmercury should be considered.

References

Duvall, S. E., and M. G. Barron. 2000. A screening level probabilistic risk assessment of mercury in Florida Everglades food webs. Ecotoxicology and Environmental Safety 47:298-305.

Hope, B. 2003. A basin-specific aquatic food web biomagnification model for estimation of mercury target levels. Environmental Toxicology and Chemistry 22:2525-2537.

Lohman, K., P. Pai, C. Seigneur, D. Mitchell, K. Heim, K. Wandland, and L. Levin. 2000. A probabilistic analysis of regional mercury impacts on wildlife. Human and Ecological Risk Assessment 6:103-130.

Moore, D. R. J., B. E. Sample, G. W. Suter, B. R. Parkhurst, and R. S. Teed. 1999. A probabilistic assessment of the effects of methylmercury and PCBs on mink and kingfishers along East Fork Poplar Creek, Oak Ridge, Tennessee, USA. Environmental Toxicology and Chemistry 18:2941-2953.

Moore, D. R. J., R. S. Teed, and G. M. Richardson. 2003. Derivation of an ambient water quality criterion for mercury: taking account of site-specific conditions. Environmental Toxicology and Chemistry 22:3069-3080.

Responses to Specific Questions Outlined in the Technical Charge

Question 1. Is this general approach reasonable? Why or why not?

The risk assessment methodology for non-fish species was based solely on exposure of methylmercury through ingestion of methylmercury contaminated prey. This approach is reasonable and appropriate for two reasons:

(1) Relative to uptake of methylmercury from aqueous sources, estimates of body normalized rates of water consumption for the species of concern and concentrations of methylmercury in surface freshwaters demonstrate that diet is the principal means of methylmercury uptake from aquatic systems for piscivorous birds and mammals.

(2) Although several field studies have documented effects of methylmercury based on organ-specific or whole-organism concentrations, there was not an indication of the concentrations of methylmercury in the diet that resulted in those concentrations that caused the effect in the organism. Consequently, for this assessment, studies of methylmercury effects on non-fish species was limited to those in which the relationship between dose (i.e., dietary concentrations and consumption) and effect was determined.

Because the 304(a) human health mercury criterion is based on fish tissue concentrations, risk for fish species based on adverse effects associated with methylmercury tissue concentrations in most appropriate.

Question 2. Are the assumptions made concerning bioaccumulation and biomagnification between trophic levels reasonable for a statewide scale model and analysis? Why or why not?

Bioaccumulation Factors (BAFs) were used to derive food chain multipliers between trophic levels (TL) 2 and 3 and TL 3 and 4. These food chain multipliers were, in turn, used (1) to calculate TL concentrations necessary to maintain the overall human dietary concentration (DC) of 0.3 mg/kg in the Average Concentration Trophic Level Approach; (2) expected concentrations in TL 2 and TL 3 if the limiting methylmercury concentration is 0.3 mg/kg in trophic level 4 fish (i.e., Highest Trophic Level Approach) and (3) predicted dietary concentrations for each of the threatened and endangered wildlife species considered under the Average Concentration Trophic Level Approach and the Highest Trophic Level Approach.

The BAFs were from draft national BAFs presented in the EPA's methylmercury criterion document (Table 1) and presented here for discussion.

Table 1. Draft BAFs for methylmercury empirically derived from field data collected across the United States and reported in the open literature. Based on BAFs calculated from lotic and lentic systems.

	BAF Trophic Level 2	BAF Trophic Level 3	BAF Trophic Level 4
5^{th} Percentile	18,000	74,000	250,000
Draft national values (approx geometric mean, 50^{th} percentile)	120,000	680,000	2,700,000
95^{th} Percentile	770,000	6,200,000	28,000,000

The document currently under review (page 5), used the approximate geometric mean BAF for each trophic level to calculate the food chain multipliers, based on (I assume) the assumption that the geometric mean best represents the wide range of BAFs from across the nation. From Table 1, and as the EPA acknowledges in the criterion document, it is evident that the range for BAFs for each trophic level varies by at least 10 (TL 2) to approximately 100 fold (TL 3 and TL 4). Moreover, the criterion document states "EPA fully recognizes that the approach taken to derive mercury BAFs collapses a very complicated non-linear process, which is affected by numerous physical, chemical, and biological factors, into a rather simplistic linear process. EPA also recognizes that uncertainty exists in applying a National BAF universally to all water bodies of the United States. Therefore, in the revised 2000 Human Health Methodology (EPA , 2000) we encourage and provide guidance for States, Territories, Authorized Tribes, and other stakeholders to derive site-specific field-measured BAFs when possible. In addition, should stakeholders believe some other type of model may better predict mercury bioaccumulation on a site-specific basis they are encouraged to use one, provided it is scientifically justifiable and clearly documented with sufficient data" (page A-18)

Using a method similar to that employed in the document under review to calculate food chain multipliers and the 5^{th} and 95^{th} percentiles for the BAFs, results in lower and upper bounds of the range of the food chain multiplier for TL 4 of at least

250,000/74,000 = 3.4

and 28,000,000/6,200,000 = 4.5.

In turn, the lower and upper bounds of the range of the food chain multiplier for TL 3 is at least

74,000/18,000 = 4.1

and 6,200,000/770,000 = 8

Although, the range of the calculated food chain multipliers is not as great as that of the BAFs from which they are derived, based on the EPA’s admission of the limitation of the draft national BAFs and the importance of the food chain multipliers to the risk assessment in question, the use of empirically derived national BAFs may not be appropriate for a statewide scale model and analysis, especially considering that BAFs from aquatic systems in California are available and could be used in the assessment.

References
U. S. EPA. 2001. Water quality criterion for the protection of human health: methylmercury. EPA-823-R-01-001. Office of Science and Technology, Office of Water, U.S. Environmental Protection Agency Washington, DC.

Question 3. Do you believe another method (or a variation of this method) might present better results? Why or why not? Stated another way: are there any elements of this analysis missing which you believe need to be included or would otherwise strengthen the analysis?

The appropriateness of use of the draft national BAFs could be evaluated by (1) calculating BAFs for lakes and streams in California and (2) comparing the draft national values to the 95% CI (for example) for the values from California.
I believe that there are several variations of the method used in the risk assessment that might present “better” results. These are listed in order of predicted increasing precision and accuracy.

(1) The document used the geometric mean national BAFs and then calculated the food chain multipliers for each trophic level. A better approach may be to calculate food chain multipliers for TL 3 and TL 4 for individual lakes and streams from the national data and then obtain the geometric mean food chain multipliers for each trophic level. The geometric mean values from the draft national BAFs for each TL probably are probably not derived from the same set of lakes and stream. This will undoubtably remove water bodies from the analysis, but will probably result in better estimates of food chain multipliers between trophic levels by using data from lakes and streams where BAFs for both trophic levels (used in calculating food chain muliplier) were actually derived.

(2) Determine statewide geometric mean BAFs from data for lentic and lotic systems in California only. California lakes and streams are included in the national data set.

(3) Same as (1), but use only lakes and streams from California.

Question 4. What are the reviewers comments concerning the use of TLRs to enhance the analysis when determining targets for wildlife species that feed from multiple trophic levels?

In evaluating the use of TLRs, it is instructive to first examine the concerns in using food chain multipliers raised by the document (USFWS 2004) evaluating methylmercury TMDLs for the Cache Creek and Sacramento-San Joaquin Delta watersheds.

The food chain multiplier was used in the document currently under review (USFWS 2003) to determine methylmercury concentrations in various trophic levels consumed by humans that, when combined, correspond to overall dietary concentrations of 0.3 mg/kg. In essence, food chain multipliers establish a linkage value for average methylmercury concentrations between trophic levels. The food chain mulipliers were derived from ratios of geometric mean BAFs for trophic levels 4, 3, and 2 from a national data base. For the Average Concentration Trophic Level Approach, these linkage values were then used in combination with data from the human health methylmercury criterion document on human consumption of fish from these trophic levels to calculate the concentrations of methylmercury in each trophic level necessary to maintain the overall dietary concentration of 0.3 mg/kg. The same linkage values were also used in the Highest Trophic Level Approach to calculate concentrations of methylmercury in trophic levels 3 and 2 if 0.3 mg/kg is the limiting concentration of trophic level 4 fish. The food chain multiplier can also be calculated from existing tissue methylmercury concentrations in fish from the various trophic levels.

USFWS (2004) on page 5, indicates that "Both of these approaches to determining food chain multipliers assume there is a direct consumption link between trophic levels, with methylmercury concentrations in the higher trophic level fish resulting from ingesting the concentrations found in fish from the next lower trophic level". I respectfully disagree. Based on how it is calculated, I suggest that the term "food chain multiplier" is a misnomer. In fact, using the ratio of BAFs results in a value that is more appropriately call a "trophic level multiplier". The linkage variable (i.e., food chain multiplier) for TL 4, for example, is based on the geometric mean BAFs for water to trophic level 4 fish and for water to trophic level 3 calculated from a national database of BAFs. No predator-prey relationship in these BAF values is inferred. As indicated in my response to question 3, I imagine that not even the same identical set of water bodies were used to calculate the BAFs for each trophic level. (I have not examined the national database.) Similarly, using average tissue concentrations in fish from different trophic levels to calculate the linkage ratio also makes no stated assumptions about predator-prey relationships. All members of the same fish species are probably assigned to trophic levels based on the generic diet of adult members of the species as fish demonstrate a fairly strong ontogenetic shift in diet—even TL 4 fish are zooplanktivorous or invertivorous at some stage of life. The average mercury concentration for the TL is undoubtably based on the average concentration of all fish, regardless of size, for the species assigned the TL designation.

However, I do agree with USFWS (2004) that it is more appropriate to derive and apply a linkage value to size classes of fish that are actually consumed by the wildlife--especially considering that the food chain multipliers are calculated from national mean BAFs. In fact, I suggest that the TLR that USFWS (2004) recommends is really a refinement of the food chain multiplier developed by USFWS (2003). The data needed to generate the TLRs for application to a statewide assessment may be difficult to obtain and, given the role that other variables have in influencing the WV and DC (see response to question 6), TLRs may not provide an additional level of refinement desired for the risk assessment. When data is available, derivation of TLRs for site-specific criteria would be preferable and is much more defensible than linkage values based on BAFs from a national data set which, as demonstrated in my response to question 3, incorporate the wide variation in water bodies from across the nation.

Question 5. Are you aware of any other relevant and appropriate studies for test doses, body weights, food ingestion rates (including dietary composition) that should be considered and included in the discussion? What are they and why are they appropriate and relevant?

I am unaware of other studies that should be considered to derive equation parameters for development of WV or DC values for the threatened and endangered species considered in this risk assessment. However, there are three recent studies by Jason Unrine and colleagues that are pertinent to the effects of methylmercury on amphibians.

In a laboratory study, Unrine and Jagoe (2004) fed larvae of southern leopard frogs (*Rana spenocephala*) aufwuchs with concentrations of inorganic and methylmercury that bracketed those observed in aufwuchs from the field and from sites reported in the literature to be contaminated with mercury from atmospheric deposition. They analyzed the metamorphs and tadpoles for inorganic and methylmercury. Inorganic Hg comprised the main source of Hg in the diets and the relative contribution of methylmercury to total mercury in the aufwuchs decreased with increasing concentration. The ratio of methylmercury to inorganic mercury in the tadpoles was greater than that of their diet and the BAF for methyl and inorganic mercury decreased with increasing concentration of mercury in the diet. A field study of mercury in frog gut contents and carcasses conducted in South Carolina wetlands (Unrine et al. 2005), confirmed their laboratory observations that amphibian larvae inhabiting areas contaminated with Hg by atmospheric deposition could be exposed to dietary mercury with concentrations as high as 1600 ng THg/ g dry weight. The majority of THg was present as inorganic Hg and the proportion of methylmercury decreased with increasing THg concentrations in the aufwuchs.

Based on their previous study of mercury bioaccumulation, Unrine et al. (2004) fed southern leopard frog larvae with aufwuchs that had mercury concentrations and speciation similar to those observed in aufwuchs from aquatic systems contaminated by atmospheric deposition. Dietary concentrations of 1500 – 3300 ng THg/g dry wt altered development and growth, and increased malformations and mortality of larvae. Methylmercury comprised 1.5-1.9% of the total mercury in the aufwuchs. They concluded that dietary mercury exposure in habitats contaminated by atmospheric deposition could adversely affect amphibian larvae.

There are few additional studies of the effects of methylmercury on fish relevant to this document. Houck and Cech (2004) found no effects of dietary methylmercury on bioenergetics of Sacramento blackfish (*Orthodon microlepidotus*) with approximately 2.5 mg/kg mercury in the muscle tissue. However, Drevnick and Sandheinrich (2003) found that dietary methylmercury at environmentally realistic concentrations suppressed estradiol and testosterone in female and male fathead minnows (*Pimephales promelas*), respectively. Suppressed levels of estradiol corresponded to altered gonad development in female fish and reduced spawning success. This study confirmed the results of a previous study by Hammerschmidt et al. (2002) that demonstrated that environmentally realistic levels of dietary methylmercury altered reproductive success of fathead minnows.

<u>References</u>

Drevnick, P.E. and Sandheinrich, M.B. 2003. Effects of dietary methylmercury on reproductive endocrinology of fathead minnows. Environ. Sci. Technol. 37: 4390-4396.

Houck,A., and J. J. Jr. Cech. 2004. Effects of dietary methylmercury on juvenile Sacramento blackfish bioenergetics. Aquatic Toxicology (69) 107-123

Unrine, J.M., and C. H. Jagoe. 2004. Dietary mercury exposure and bioaccumulation in southern leopard frog (*Rana sphenocephala*) larvae. Environmental Toxicology and Chemistry 23: 2956–2963.

Unrine, J.M., C. H. Jagoe, W. A. Hopkins, and H. A. Brant. 2004. Adverse effects of ecologically relevant dietary mercury exposure in southern leopard frog (*Rana sphenocephala*) larvae. Environmental Toxicology and Chemistry 23: 2964–2970.

Unrine, J. M., C.H. Jagoe, A.C. Brinton, H.A. Brant, and N.T. Garvin. 2005. Dietary mercury exposure and bioaccumulation in amphibian larvae inhabiting Carolina bay wetlands. Environmental Pollution. In Press.

Question 6. Are the conclusions reached reasonable based on the information considered? Why or why not?

In general, the conclusions reached are reasonable based on the risk assessment method employed and the data used to derive the RfD, FIR, and DC. If anything, the derived WVs may be underprotective because of the low UFs used to derive the RfD for each species. The range of the maximum cumulative UF (i.e., product of the individual UFs) used in determining the species RfDs was 3 -9. A review by Duke and Taggart (2000) of the UFs used in 24 screening ecological risk assessments in California reported that the maximum cumulative UF ranged from 10 to 3000. However, they noted a "paucity of agency guidelines for selecting UF for HQ *(hazard quotient*—my note*)* calculation and little information intended for agency use on appropriate UFs for varying conditions and extrapolations". In the current risk assessment, guidance on determining the appropriate values for each uncertainty factor was obtained from USEPA (1995) and Abt Associates (1995—same as Abt Associates reference below?) but these documents also give relatively minimal guidance for value derivation. For example, with regards to the interspecies uncertainty factor (UF_A), USEPA (1995) indicates that information on species sensitivity differences "provide support for the 1-100 recommended range for the value" but stress that "the actual selection of an interspecies UF for application to a particular situation must be made on a case-by-case basis and requires the use of *best professional judgement"* (page 20; emphasis mine). Abt Associates (1995) suggested that an interspecies uncertainty factor of 10 was justified (as cited in Chapman et al. 1998—I did not have access to Abt Associates 1995). Calabrese and Baldwin (1993) recommended an UF_A of at least 10 to protect endangered species.

As a further illustration, if an interspecies uncertainty factor of 10 is used, then the WV for the sea otter becomes 0.0055 mg/kg diet and the calculated DC for the Average Concentration Trophic Level Approach exceeds the WV by 10 fold and by more than 4 fold for the Highest Trophic Level Approach. Under this scenario, the EPA's human health TRC (0.3 mg/kg) is likely to result in a dietary exposure that would place sea otters at risk, contrary to the conclusion reached by the document under review (page 64). A similar analysis could be conducted for the other species. The document states that the interspecies uncertainty factor of 1 was selected because the southern sea otter is in the same taxonomic family (although different genera) as the mink from which the TD was determined and, therefore, should have similar sensitivity. However, if protection of individuals is the goal, then a careful assessment of intra- and interspecies extrapolations is warranted. For example, in the review by Chapman et al. (1998), they noted that the U.S. EPA commissioned a report by Abt Associates (1995) to support development of wildlife criteria for the Great Lakes water quality initiative. A review and analysis by Abt Associates (1995) of subchronic and chronic data from 174 toxicity studies with birds and mammals showed that 94% of the species sensitivities were within a factor of 50 of each other and *ranged 1 to 1000* (emphasis mine). Moreover, Chapman et al (1998) reported that, based on intraspecific variation in response by laboratory rodents in toxicity tests, human variability in toxicological assessments is assigned an uncertainty factor of 10 and "many ecotoxicologists assume that wildlife species tested for toxicity responses would have a greater intraspecific variation than inbred laboratory rodent strains because of greater genetic diversity". Hence, an interspecies uncertainty factor of 10 is not unreasonable.

My point here is not necessarily to criticize the values selected for each UF, nor advocate the use of a particular value. Rather, my intent is to highlight the fact, as others have more eloquently done, that the use of uncertainty factors (aka safety factors) makes some assumptions with unknown scientific justification and, hence, is perhaps one of the weakest links in an ecological risk assessment that uses hazard quotients.

For the current document under review, I concur with several of the recommendations made by Chapman et al. (1998) for the use of safety factors. Specifically, "(1) Extrapolation requires context. Any use of safety factors should be based on existing scientific knowledge and should include appropriate caveats; (2) Extrapolation is not fact. Safety factors should be used only for screening, not as threshold or absolute values; (3) Extrapolation is uncertain. Safety factors should encompass a range, not a single number". In addition, it may be appropriate to provide some indication of how the UFs used in this document compare to those used by the state for ecological risk assessment.

References

Abt Associates. 1995. Technical basis for recommended ranges of uncertainty factors used in deriving wildlife criteria for the Great Lakes water quality initiative. Final Report. Office of Water, U. S. Environmental Protection Agency, Washington, DC.

Calabrase, E. J., and L. A. Baldwin. 1993. Performing ecological risk assessments. Lewis, Chelsea, MI, USA.

Chapman, P. M., A. Fairbrother, and D. Brown. 1998. A critical evaluation of safety (uncertainty) factors for ecological risk assessment. Environmental Toxicology and Chemistry 17:99-108.

Duke, L. D., and M. Taggart. 2000. Uncertainty factors in screening ecological risk assessments. Environmental Toxicology and Chemistry 19:1668-1680.

USEPA. 1995. Great Lakes Water Quality Initiative Technical Support Document for Wildlife Criteria. EPA-820-B-95-009. Office of Water. Washington, DC

Question 7. Is this approach reasonable? Why or why not? Are there any elements of the analysis missing which you believe need to be included or would otherwise strengthen the analysis? If so, please explain.

The dual approach of using DCs based on the Average Concentration TL Approach and Highest TL Approach is reasonable and allows comparison of potential risk under the method used to derive the human health criterion (Average Concentration TL) and under the method by which most states would, I assume, apply the human health criterion (Highest TL Approach).

Question 8. What is/are the weakest part(s) of the analysis and why? Do have any suggestions for alternatives? Please discuss any other general or specific concerns you might have regarding any aspects of the analysis.

I think the analysis has shortcomings in several areas.

(1) Values for different model parameters are presented without careful statement and consideration of the assumptions that were made in deriving those parameters. This leaves the uninformed reader with the impression that the parameter values are deterministic and without variability (e.g., food chain multipliers).

(2) The UFs have a major role in determining whether or not the DCs will exceed the WVs and result in a conclusion of probable harm. The UFs values selected need stronger scientific justification other than that they were the same as those used in the Great Lake initiative and Mercury Study Report to Congress. The degree of conservatism of the risk assessment is unknown.

(3) The conclusions based on the risk assessment are deterministic without measure of the uncertainty or variability associated with model parameters or estimates of the probability of adverse effects to the species of concern. A probabilistic risk assessment model would better incorporate the uncertainty and stochasticity of model parameters and provide decision makers with better information on the probability of adverse effects to the threatened and endangered species.

PEER REVIEWER COMMENTS

DR. EDWARD SWAIN

Peer Review of “Evaluation of the Clean Water Act Section 304(a) Human Health Criterion for Methylmercury: Protectiveness for Threatened and Endangered Wildlife in California”

March 2005

By Edward Swain

Question 1. Is this general approach reasonable? Why or why not?

Yes, this general approach is reasonable. The document acknowledges challenges and limitations, and deals with them in a thoughtful and robust manner. However, it was not an effortless application of the approach. The relevant toxicological studies were read carefully and in some cases the original authors were contacted for specific information that was not in the published results. Final judgment on the appropriate uncertainty factors for several of the endangered species required application of deep expertise on the part of the authors, and in the end, the decision was sometimes based on expert opinion. When this was necessary, the rationale for this expert opinion was well documented and presented

Question 2. Are the assumptions made concerning bioaccumulation and biomagnification between trophic levels reasonable for a statewide scale model and analysis? Why or why not?

The assumptions are reasonable for assessing the potential for harm from methylmercury exposure, but not for assessing the exact effect on a given species. The document did not attempt to do the latter, so the modeling effort was appropriate. Even though there is a great deal of data represented in this document, and many calculations are taken to three decimal places, in actuality this is a coarse assessment of potential harm. I do not see evidence that the assessment procedure is consistently tilted toward conservative assumptions, which might predict harm when none would exist. Rather, I see that this approach is utilizing average conditions. For instance, food chain multipliers were calculated from the draft national BAFs presented in EPA’s 2001 methylmercury criterion document (discussed in the text on page 5, but not listed in the references for some reason). These national BAFs are averages, but the 90% ranges in the BAFs have at least an order of magnitude of variation (p. A-14 of the 2001 document). The text on page 5 goes on to state that if site-specific BAF data are available, they could be used. I think that the point is that the use of average BAF data is reasonable unless good site-specific BAF data are available. But good site-specific BAF data are rarely available, and the potential for harm can be assessed using average data on a statewide scale.

Question 3. Do you believe another method (or variation of this method) might present better results? Why or why not? Stated another way: are there any elements of this analysis missing which you believe need to be included or would otherwise strengthen the analysis?

Better results would of course be produced if site-specific data were available—data associated with the species of interest. But, in the absence of site-specific data, the RWQCB's approach described in USFWS (2004) document is a desirable refinement. However, I feel that the distinctions made between the food chain multiplier (FCM) and trophic level ratio (TLR) is largely semantic, since the FCM is based on empirical trophic level data. The important point made in USFWS (2004) is the attention paid to the size of fish food items, since mercury concentration is very dependent on the size of a particular fish within a species. If there were any one change that could be made to the method that would be valuable, it would be adjusting the concentration of prey fish in a particular trophic level for the size of those fish. Some might argue that this is too fine a point, given the large variation already acknowledge within a trophic level. But, if a species is eating 100% of a small size fish in TL4, it could easily make a 3-fold difference in assumed concentration, equivalent to a jump in the uncertainty factor. 3-fold changes might not make much difference for some contaminants, especially completely anthropogenic chemicals such as PCBs, but for a chemical with significant natural levels that is perhaps only a factor of 3 lower than current ambient concentrations, we cannot afford to be cavalier with 3-fold differences. In fact, only one species in this analysis has a DC to WV ratio of appreciably more than three (that is, 300%)—the Yuma clapper rail (962%), and then only when the UF_A is 3, rather than 1 (Table 5). The Light-footed clapper rail ratio is 408% with a UFA of 3, but in the grand scheme of things, 400% is pretty similar to 300%. It is noted on page 81 that specific research is needed to determine exactly how vulnerable the light-footed clapper rail is. The document is judicious in its call for research to address weaknesses in the analysis.

Question 4. What are the reviewers' comments concerning the use of TLRs to enhance the analysis when determining targets for wildlife species that feed from multiple trophic levels?

The use of TLRs is desirable if it is a means to more realistically model the size of prey that are actually being consumed. I have already discussed this issue in my response to Question 3.

Question 5. Are you aware of any other relevant and appropriate studies for test doses, body weights, food ingestion rates (including dietary composition) that should be considered and included in the discussion? What are they and why are they appropriate and relevant?

No, I am not aware of any relevant studies that should have been considered for this analysis. The breadth of review and the thoughtful interpretation impressed me.

Question 6. Are the conclusions reached reasonable based on the information considered? Why or why not?

The conclusions are reasonable, based on the information considered. The information is limited in many ways, and so it is appropriate to not over-interpret it by creating a misleading degree of detail. This analysis certainly reveals which species are more at risk, and therefore which species for which more data should be obtained.

I found the premise of the document a bit inefficient, from a scientist's point of view. I had to keep remembering that the endangered species were not being evaluated against present ambient conditions, but, rather, what if the mercury concentrations of fish were in compliance with the water quality criterion of 0.3 mg/kg of edible portions. And, then, to further complicate the scientific analysis, there were two different approaches within which this criterion might be met. I felt that this series of premises made the document difficult to read and unnecessarily difficult to determine if the conclusions were reasonable. However, after multiple readings, I am confident that I have followed the logic and find it reasonable. In fact, having been a peer reviewer of many scientific documents, I find this document uncommonly well written. I feel that the authors were given tight constraints within which to work, and that an admirable job was accomplished. A much more satisfying product would have been produced had site-specific information been made available to the authors, but such information was not available.

Question 7. Is this approach reasonable? Why or why not? Are there any elements of the analysis missing which you believe need to be included or would otherwise strengthen the analysis? If so, please explain.

Some might suggest that the analysis would be strengthened through the use of probabilistic risk assessment methods. This suggestion is attractive, in that it would yield quantitative estimates of the proportion of individuals in the population that are at what degree of risk. Such estimates would, however, give a false sense of knowledge, given the nature of the inputs to this analysis. A probabilistic assessment would be more appropriate with site-specific information, in contrast to the largely national average data used in this analysis. I am not completely happy with the exact numbers that are produced with the present approach, but the discussion that accompanies the exact numbers utilizes them appropriately—they are interpreted more qualitatively than quantitatively. Because this document presented two different approaches, the "Average Concentration Trophic Level Approach" and the "Highest Trophic Level Approach", and always discussed both results, there was a degree of temperance to the conclusions, such as in IX.B. (pages 81 to 82).

Question 8. What is/are the weakest part(s) of the analysis and why? Do you have any suggestions for alternatives? Please discuss any other general or specific concerns you might have regarding any aspect of the analysis.

The weakest parts of the analysis are largely issues that the authors cannot control: Use of average BAFs because site-specific data was not collected; Unknown degree of methylmercury sensitivity of an untested species; overly broad dose spacing of the toxicity tests that were performed (NOAEL-LOAEL issues); Uncertainty in converting subchronic data to chronic. These weaknesses were acknowledged and adequately dealt with in the document, but certainly are still weaknesses. The USFWS (2004) document discusses a weakness that could conceivably be adequately dealt with even without a great deal of new data: within a trophic level, fish accumulate mercury in a predictable manner as they grow. To incorporate this knowledge in modeling the mercury concentration of food would increase accuracy of the modeling effort. If predators are restricted to a given size of fish within a trophic level, it would increase the accuracy of mercury exposure estimates to include that information in some quantitative manner.

References Cited

USEPA. 2001. Water Quality Criterion for the Protection of Human Health: Methylmercury. EPA-823-R-01-001. January 2001

USFWS. 2004. Evaluation of Numeric Wildlife Targets for Methylmercury in the Development of Total Maximum Daily Loads for the Cache Creek and Sacramento-San Joaquin Delta Watersheds. 30 pp.

PEER REVIEWER COMMENTS

DR. MARTI WOLFE

Evaluation of the Clean Water Act Section [CWA] 304(a) Human Health Criterion for Methylmercury: Protectiveness for Threatened and Endangered [T&E] Wildlife in California"

Peer Review Summary
Marti F. Wolfe

Question 1. Is this general approach reasonable? Why or why not?

My overall assessment is that this approach is mostly reasonable, well researched and well documented, *given the current state of knowledge*. Therefore my comments are mostly address refinements, and/or additional specificity that can be used to strengthen the approach as results from studies currently being conducted become available.

Methodology - Bioaccumulation/Biomagnification:

The analysis assumes that although aquatic food webs vary considerably across all waters of California, the aquatic organisms can generally be classified into four trophic level categories.

> This is reasonable as long as seasonal variation, migrational and/or opportunistic variations in trophic status are taken into consideration.

The analysis further assumes that biomagnification of methylmercury among trophic levels is sufficiently consistent across waterbodies that food chain multipliers can be estimated for this analysis based on draft national bioaccumulation factors (BAFs). If more site-specific BAF data exist, then that data can (and should) be used in place of the draft national BAFs to determine better site specific multipliers.

> Draft national bioaccumulation factors are useful for calculations where site-specific data are lacking, but local derived factors are preferable. These can vary considerably even within a relatively homogeneous water system (Hope 2003).
>
> Where more site-specific data do not exist, they should be sought, especially in habitats/systems where mercury behavior is highly variable and/or those impinging on threatened and endangered species, such as the Bay-Delta system.
>
> Important work on trophic transfer on Bay and Delta marshlands has recently been completely by Letitia Grenier at UC Berkeley, and will soon be available in citable form.

Question 2. Are the assumptions made concerning bioaccumulation and biomagnification between trophic levels reasonable for a statewide scale model and analysis? Why or why not?

Variation in California due to differences in Hg inputs, temperature, salinity and in methylation regimes, as mentioned above, can result in variability in biomagnification and bioaccumulation. As recommended by Weiner et al, habitats areas and trophic pathways associated with bioaccumulation and biomagnification of methylmercury to elevated concentrations should be identified (Weiner et al. 2003).

Furthermore, several mercury-bearing wetland areas of California are scheduled for restoration. This will provide much-need additional habitat for T & E species, but it is recognized that these activities may result in temporary elevations or 'pulses' of available and reactive mercury into food chains (Slotton et al. 2001). These areas will need to be monitored to determine if TRCs calculated pre-restoration are adequately protective during transient perturbations (Caldwell 2004).

Question 3. Do you believe another method (or a variation of this method) might present better results? Why or Why not? Stated another way: are there any elements of this analysis missing which you believe need to be included or would otherwise strengthen the analysis?

The State of California's Regional Water Quality Control Board used a similar analysis to determine numeric targets to protect wildlife in the Cache Creek and the Sacramento-San Joaquin Delta watersheds. Instead of determining fish tissue targets based on trophic levels, the Board developed targets based on specific size classes and species of fish that are believed to be consumed by the specific species of concern. That is, instead of deriving simple food chain multipliers based on the biomagnification expected when fish from one trophic level consume prey from the next lower trophic level, the Board derived trophic level ratios (TLRs) based on concentrations in fish prey that are similar in size (no-predator-prey interaction) but occupy different trophic levels. This discussion/explanation is found at pages 5 and 6 of the document entitled, "Evaluation of Numeric Wildlife Targets for Mercury in the Development of Total Maximum Daily Loads for the Cache Creek and Sacramento-San Joaquin Delta Watersheds".

> The Average Concentration Trophic level and Highest Trophic Level comparison approach is more sensitive, if more demanding to apply, the Regional Board's
>
> One particular strength of this work compared to other comparable efforts is that fish are regarded as *targets* of toxicity, and not only *vectors* of toxicity; and a summary of methylmercury toxicity to fish is included.
>
> Work in progress will provide improved data on the impacts of MeHg to sturgeon (white and green), Sacramento splittail, and Delta smelt. These should be incorporated as they become available.
>
> Invertebrates, both as targets and vectors, are under-represented. They should be included because of their abundance in the food web, and because the impact of mercury in invertebrates is much less well studied. For example we can use a rule of thumb for vertebrates that 90 – 99 % of total mercury in tissues is methylmercury, but the same blanket metric cannot be employed for inverts. This could be important in calculating the dose of animals that feed heavily on inverts.

Question 4. What are the reviewers' comments concerning the use of TLRs to enhance the analysis when determining targets for wildlife species that feed from multiple trophic levels?
"

Methodology - Risk Assessment:
WV. The calculations employed in this work are consistent with current methods for arriving at Wildlife Values. The method has undergone some evolution since introduced in the GLI, and may continue to do so as improved data for its input values becomes available.

Test Doses:

Mammalian

Although the 30-year-old results of Wobesor were the best available when this document was written, newer studies on mink have recently been completed at the Nova Scotia Agricultural College using ecologically relevant concentrations of Hg (Nil Basu, pers. comm., McGill University, Montreal, Canada), and multiple sensitive endpoints (Basu et al. 2005, Basu 2005, Stamler et al. 2005). These data should be incorporated as they become available.

Avian

Work to refine and strengthen the avian WV is in progress, but less close to availability than the efforts for mink. Kenow, USGS, Meyer, Wisconsin Department of Environmental Quality and co-workers, will undertake an egg injection study in loons based on the methods of Heinz, in 2005 and 2006. Frederick and Spaulding, University of Florida, have currently under way a feeding study using captive great egrets. A WV using common loons has recently been presented by Evers and co-workers at Biodiversity Research Institute (Evers et al. 2003). Concurrent efforts of Schwarzbach and Suchanek will supply much-needed specifics from the Bay-Delta system; these will further enhance the precision of the model.

Uncertainty Analyses: The treatment of uncertainty analysis in this document reflects the most recent peer-reviewed sources for methylmercury, and is treated with thoroughness. However, recent work from other than wildlife species and substances other than mercury may be consulted to some benefit (Gray et al. 1998, Irwin et al. 2004, Lock and Smith 2003, Orphanides 2003, Snape et al. 2004, Waters and Fostel 2004).

Body Weights:

For T & E species, the level of concern may need to shift from population-level to individual animal effects, in which case not only body weights, different metabolic responses of juveniles may need to be incorporated (Stark et al. 2004).

Food Ingestion Rates: Use of the Nagy reference represents the current standard. Further refinements may be obtained by consulted more recent work (McKechnie and Wolf 2004, White and Seymour 2004).

Question 5. Are you aware of any other relevant and appropriate studies for test doses, body

weights, food ingestion rates (including dietary composition) that should be considered and included in the discussion?

I have included these under earlier questions.

Question 6. Are the conclusions reached reasonable based on the information considered? Why or why not?

Yes, the conclusions reached are reasonable, and the document as it stands is suitable for regulatory decision making. However, as results from studies currently underway become available, they should be incorporated – that is, the document should be considered as an evolving method, not a fixed and unchangeable entity.

Question 7. Is this approach reasonable? Why or Why not? Are there any elements of the analysis missing which you believe need to be included or would otherwise strengthen the analysis? If so, please explain.

There is insufficient attention given to data coming from outside the conventionally accepted feeding-study-in-a-wildlife species. Once again, this is less a reflection on the authors of this documents than a failure of vision of all of us in the field. Toxicogenomics, QSAR (quantitative structure activity relationships) inter-species toxicokinetics extrapolations, etc in the human health and laboratory animal arena have moved far beyond us in wildlife. Our failure to close this gap only deprives us of powerful tools which we can most efficaciously employ in risk assessment for T & E species but will benefit wildlife species in general.

Question 8. What is/are the weakest part(s) of the analysis and why? Do you have any suggestions for alternatives? Please discuss any other general or specific concerns you might have regarding any aspect of the analysis.

The weakest part of the analysis is that its most core component, the WV, is based on data ~30 years old. This is no reflection on the quality of the effort; it is simply the state of the science, one that everyone working in the field laments. The work in progress that I have described and/or referenced above will do much to add refinements and enhance predictability of MeHg risk assessment to wildlife species.

The in-progress mammal data promise to be more robust and more immediately applicable by virtue of coming from a single lab, as a coordinated effort.

This was also the case with the Heinz-Hoffman mallard studies of the 1970-s and is one reason those data have stood the test of time and continue to be useful for current risk assessments.

The avian work currently in progress, in contrast, represents the effort of a number of investigators in government, university, and NGO organizations. Although the quality of the data is anticipated to be high, there is reason for concern that it may not be a further refinement over Heinz-Hoffman because of differences in protocol, endpoints measured, etc.

I propose that a coordinating council for avian MeHg be established, that asks the following questions:

1) What data do we need besides Heinz-Hoffman for an improved avian MeHg WV?
2) Which of those data can be anticipated from studies currently under way?
3) What additional studies (species, endpoints) are needed to fill the gaps identified by the answer to #2?

Until those data are obtained and validated via the peer-review process, this document represents the best that we have. It is a very commendable work.

Basu, N., K. Klenavic, M. Gamberg, M. O'Brien, R. D. Evans, A. M. Scheuhammer and H. M. Chan. 2005. Effects of mercury on neurochemical receptor binding characteristics in wild mink. Environmental Toxicology and Chemistry 24: 39-45.

Basu, N., Stamler, C.J., Loua, K.M., and Chan, H.M. 2005. 2005. An inter-species comparison of mercury inhibition on muscarinic acetylcholine receptor binding in the cerebral cortex and cerebellum. Toxicology and Applied Pharmacology In press:
doi:10.1016/j.taap.2004.09.009:

Caldwell, B. a. 2004. Mercury Technical Memorandum, South Bay SaltPond Restoration Project. San Francisco, CA. California State Coastal Conservancy U.S. Fish and Wildlife Service, California Department of Fish and Game.

Evers, D. C., O. P. Lane and L. Savoy. 2003. Assessing the impacts of methylmercury on piscivorous wildlife using a wildlife criterion value based on the Common Loon, 1998-2002. BioDiversity Research Institute, Falmouth. Maine.

Gray, L. E. J., J. Ostby, C. Wolf, C. Lambright and W. Kelce. 1998. The value of mechanistic studies in laboratory animals for the prediction of reproductive effects in wildlife: Endocrine effects on mammalian sexual differentiation. Environmental Toxicology And Chemistry 17: 109-118.

Hope, B. 2003. A basin-specific aquatic food web biomagnification model for estimation of mercury target levels. Environ Toxicol Chem 22: 2525-37.

Irwin, R. D., G. A. Boorman, M. L. Cunningham, A. N. Heinloth, D. E. Malarkey and R. S. Paules. 2004. Application of toxicogenomics to toxicology: basic concepts in the analysis of microarray data. Toxicol Pathol 32 Suppl 1: 72-83.

Lock, E. A. and L. L. Smith. 2003. The role of mode of action studies in extrapolating to human risks in toxicology. Toxicol Lett 140-141: 317-22.

McKechnie, A. E. and B. O. Wolf. 2004. The allometry of avian basal metabolic rate: good predictions need good data. Physiol Biochem Zool 77: 502-21.

Orphanides, G. 2003. Toxicogenomics: challenges and opportunities. Toxicol Lett 140-141: 145-8.

Slotton, D. G., S. M. Ayers, T. Suchanek, R. D. Weyand, A. M. Liston, C. Asher and D. C. Nelson. 2001. The Effects of Wetlands Restoration on the Production and Bioaccumulation of Methyl Mercury in The Sacramento-San Joaquin Delta, California. CALFED Proposal

Snape, J. R., S. J. Maund, D. B. Pickford and T. H. Hutchinson. 2004. Ecotoxicogenomics: the challenge of integrating genomics into aquatic and terrestrial ecotoxicology. Aquat Toxicol 67: 143-54.

Stamler, C. J., N. Basu and H. M. Chan. 2005. Biochemical markers of neurotoxicity in wildlife and human populations: Considerations for method development. Journal of Toxicology and Environmental Health In press; accepted: Dec 17, 2004:

Stark, J. D., J. E. Banks and R. Vargas. 2004. How risky is risk assessment: the role that life history strategies play in susceptibility of species to stress. Proc Natl Acad Sci U S A 101: 732-6.

Waters, M. D. and J. M. Fostel. 2004. Toxicogenomics and systems toxicology: aims and prospects. Nat Rev Genet 5: 936-48.

Weiner, J. G., C. G. Gilmour and D. P. Krabbenhoft. 2003. Mercury Strategy for the Bay-Delta Ecosystem: A Unifying Framework for Science, Adaptive Management, and Ecological Restoration. California Bay Delta Authority

White, C. R. and R. S. Seymour. 2004. Does basal metabolic rate contain a useful signal? Mammalian BMR allometry and correlations with a selection of physiological, ecological, and life-history variables. Physiol Biochem Zool 77: 929-41.