

Final Comments of Tim Thompson on the Ecological Risk Assessment for the Housatonic River.

This memo represents my final comments to the ERA conducted for the Housatonic River as part of the Peer Review Panel Process. This document is modified from the pre-meeting comments submitted January 5, 2004, based upon the review and discussion from the Peer Review Panel meeting held January 14 – 16.

The format here follows that prescribed in the Charge to the Peer Reviewers provided by SRA International Inc. I have deviated somewhat from that format by answering each of the Charge Questions as they relate to the Assessment Endpoints identified as part of Charge Question 3.

I would like to reiterate what I have said at our Peer Review meetings, and in my January 5 comments – this is an amazing document with a considerable amount of thought and effort expended by both EPA and GE scientists and ecologists. I extend my complements to each of them for the tremendous effort – both in terms of the science they practiced and the Herculean effort needed to bring these documents together. I also reiterate that it has been my pleasure to observe the very professional way in which EPA and GE interact on this project. Reasonable people can, and will disagree. It is very nice to observe the process do so in a reasonable way.

An explanation to the way I approached my comments to the ERA may be helpful. I agree strongly with the following statement made in Section 2.8.1 (page 2-59): “Ultimately, the value of an ERA depends on whether it can be used to determine if a baseline risk is present and to support managerial decisions.” As such, I have tried to “truth test” as many of the conclusions as was possible within the time frame allotted. This meant evaluating in detail the Supplemental Investigation Work Plan for the ERA, requesting the same – and receiving work plans from GE (thank you!), reading through the RCRA Facilities Investigation, as well as the ERA, supporting Appendices, and all of the other supporting documents. Where I have made comments, it is in the interest of pointing out what was apparently intended (work plans), what was actually done (appendices and supporting documents), and how that was used in the characterization of risk.

I would like to acknowledge the remarkable diversity and intellect of my Peer Review colleagues. This was truly a powerful assemblage of scientist who critically, and objectively evaluated the studies conducted by both EPA and GE with no bias – save for what their experience and intellect told them was right and wrong. I was humbled to be considered amongst such a group, and believe that the ERA can only be strengthened by their individual, and our collective input.

Benthic Infaunal Endpoint

Charge Question #1

“Was the ecosystem of the Housatonic River watershed properly characterized, and was this information appropriately applied in the Problem Formulation and subsequently in the ERA?”

There is a tremendous amount of ecosystem information that has been generated that could be used to help articulate the importance of aquatic invertebrates to the Housatonic River. Numerous studies have been conducted related to aquatic insects with the Housatonic River, and in the PSA. Given that the basis for evaluating aquatic invertebrates is the importance in the food chain and transfer of PCBs into higher organisms, a discussion of the higher trophic level organisms and the prey upon which they depend should be better articulated for this assessment endpoint.

One consideration is to make more use of the riverine classifications in describing the various benthic habitats, expected species, and anticipated routes of exposure. Within the Natural Communities descriptions, for medium gradient streams the soil type is classified as cobble, gravel and sand, with moderate stream flows. Within the classification it would then have been useful to list the type of aquatic invertebrates expected in those habitats – and potentially have actually quantified the species existing there (presumably higher in epibenthic organisms, and in particular EPT species). Aquatic invertebrates in this habitat would be less likely to be exposed through fine-sediment contact and porewater uptake, and would receive PCBs through the water column and/or contaminated food. Within depositional pools, such as the moderately alkaline pond, more benthic infauna would be encountered, and would include a significant component of the invertebrate biomass exposed to PCBs through sediment and porewater ingestion. But in addition, given the high density of emergent and floating vegetation, a high degree of “benthic” invertebrate production would also be expected from within the plant mass.

A general comment here, and throughout the ERA, is that figures showing spatial variation of sample sites, including sediments, soils, and tissues should be included to help illustrate the depth and breadth of coverage.

The ERA also needs to make a better explanation of the types of communities encountered, relative to the grain size and organic carbon contents in those sediments. As noted by Dr. Oris in his pre-meeting comments [3.1(c).1], the relationships between chironomid abundance, taxonomic diversity relative to substrate (fine vs. coarse) potentially obfuscate any relationship between tPCBs and community structure. Without a complete discussion of communities, substrates, riverine conditions, and organic carbon, the relationship to tPCBs and toxicity is tenuous.

It would have also been appropriate to have conducted community analyses throughout the river, making use of kick nets and drift nets in the riffle-runs, emergent insect traps during late Spring and summer months to assess potential prey species (hence exposure) for insectivorous birds and mammals, and potentially gut-content analyses for fish species in the River. It would appear that at least a part of this information does exist, and to the degree practical it would enhance the ERA to summarize this into a qualitative description of the benthic ecosystem. It would also be useful to include in a general discussion any findings related to crayfish, mussels, and dragonflies to the overall description of the benthic communities.

Moving onto the actual assessment and measurement endpoints, the stated assessment endpoint was listed as “community structure, survival, growth and reproduction of benthic invertebrates”. Benthic invertebrates, by definition, would be animals living on or near the bottom of an aquatic habitat (Thorp and Covich, 1991). This is fairly broad and includes organisms living within the substrate (infauna) or upon the substrate (epifauna). A number of those same species will live in or upon aquatic vegetation. It appears that the ERA chose to focus on a narrow subset – principally benthic infauna. This is not only reflected in the overall sampling design, but is also captured in EPA’s response to the specific question – define benthic invertebrates (Question TT32 – EPA’s Response to Questions dated 12/11/03). EPA’s response was that

“...benthic invertebrates refer to invertebrates that are exposed via sediment and porewater uptake pathways. These organisms include “infauna” and “epifauna” but exclude “water column invertebrates””

This is an important distinction – and should be reflected within the assessment endpoint. Exposure is “via sediment and porewater”, and I would submit then by definition the assessment is *community structure, survival, growth and reproduction of benthic infauna*.

In summary, and germane to my subsequent arguments below, (1) the selection of measurement endpoint should be consistent with the route of exposure stated (sediment and porewater), (2) risks to a very important source of aquatic insect prey for fish, birds and insectivorous mammals within the PSA is not addressed by this assessment endpoint, and (3) management actions that could come from this endpoint are only applicable to those areas where soft sediments and benthic infauna exist. In fact, the inconsistency within the assessment and measurement endpoints, along with some specific data quality problems, impair the use of these data in formulating a risk-management decision.

Charge Question #2

“Was the screening of contaminants of potential concern (COPCs), selection of assessment and measurement endpoints, and the study designs for these endpoints appropriate under the evaluation criteria?”

The use of sediment quality thresholds and water quality criteria for screening COPCs are very well established in the scientific and ecological risk assessment community. Methods used by the ERA for screening are appropriate.

As presented above, the ERA's intent was to assess risks to infaunal organisms – not to aquatic invertebrates in general. That the ERA's intended is to assess benthic infauna is reflected within the Supplemental Investigation Work Plan (Weston 2000). Section 7.2.4 poses the question “are ambient PCB levels in sediment sufficient to cause adverse effects in benthic organisms..”. Within Appendix A.13, it states that “Sampling of benthic macroinvertebrates inhabiting soft (depositional) sediments will be conducted at 13 locations in the Housatonic River watershed”.

What is very clear and important is that the ERA never intended to assess overall benthic infaunal community health – rather it set out to see if a dose/response relationship could be established between levels of PCBs and the specific measurement endpoints. Section 2 of Appendix A.13 to the Work Plan states that

“The benthic macroinvertebrate investigation is sharply focused on providing information specific to the three data quality objectives outlined above, rather than providing a full characterization of the benthic macroinvertebrate community, which is beyond the scope of this study.” (emphasis added)

Those objectives were defined as evaluating PCBs, other contaminants and effluent from the WWTP, providing tissue concentrations of benthic macroinvertebrate groups for use in the modeling, provide “one of the three components of the Sediment Quality Triad approach to evaluating the potential ecological risk across a range of PCB concentrations in the sediment.” Further support that measurement endpoints were selected not to assess overall community health is found in Section 3.2 of Appendix A.13 which states that

“The rationale for selection of the 13 locations to be sampled in the benthic macroinvertebrate study is presented in Subsection 2.1.2. The locations are not intended to be representative of the entire river but rather are intended to encompass the range of sediment PCB concentrations in the Lower River between the Confluence and Woods Pond..”

“Of the remaining five target stations, four will be selected to (1) provide data from locations with PCB concentrations that expand upon the range of concentrations represented at the Triad locations and (2) investigate the potential effects of the WWTP discharge on benthic communities in the river.”

This argument is crucial to supporting the assessment – if a dose/response relationship cannot be demonstrated within the measurement endpoints, then there is very little utility in the acquired data. Based upon my read of this information, the assessment endpoint

would be reformulated into the following testable hypothesis: *There is a dose-dependent response to PCBs in community structure, survival, growth and reproduction of benthic infauna.* That is the way the study is currently constructed, so this would be a more appropriate assessment endpoint.

Having noted that, I am in agreement with some of my Panel colleagues who noted that the Assessment Endpoint should be the long-term persistence of benthic invertebrates. Dr. Forbes' comment that "...consideration be given to redefining the assessment endpoints: reproduction, growth, and survival as measurement endpoints for the target species considered and that the assessment endpoints be redefined as impairments to long-term persistence of populations of target taxa". Taking this approach would, however, require more of a reliance on community structure than the bioassay results.

I take the dissenting position from EPA and some of my Panel colleagues in that I do not agree that measurement endpoints centered on benthic exposure represent a "worst case exposure", which would be representative of all infaunal and epifaunal species. The risk assessment is about important receptor groups and the appropriate receptor pathways. To state that results of chironomids bioassays are indicative of protection to EPT or odonates epifaunal species, or to state that daphnid bioassays are applicable to benthic infauna, does not provide a direct link between exposure, effect, and importance to the food web on the Housatonic River. On that note, the Panel was generally in agreement that daphnids were not appropriate measurement endpoints to the stated assessment endpoint.

A final note, it is interesting to note that the list of assessment and measurement endpoints between the Work Plan (Section 7) and the ERA are different. Table 2.8-1 adds comparison of tissue chemistry to tissue effects thresholds, and drops survival and physiology of freshwater mussels. An explanation for this should be included in the final ERA. Tissue residue values will be discussed more below.

Charge Question #3

"For each of the 8 assessment endpoints evaluated in the ERA (listed in Attachment B, and for which a specific Section and Appendix was prepared), address the following questions (discuss and label responses as 3.(assessment endpoint number).(question letter) for consistency):"

"(a) Were the EPA studies and analyses performed (e.g., field studies, site-specific toxicity studies, comparison of exposure and effects) appropriate under the evaluation criteria, and based on accepted scientific practices?"

Within the context of defining risks to PCBs to benthic infauna based upon a dose/response relationship in the legs of the Sediment Quality Triad, the planned field studies, toxicity studies, and accompanying measurements were all accepted scientific

practices. There are elements within each of the “legs” that influence the quality and use of the findings, are discussed below.

Sediment Measurements of tPCBs The variability and range of tPCB levels in sediment are perhaps the most greatest impediment to the utility of these data for the assessment and characterization of risk. The high spatial and temporal variability in sediment PCB concentrations created difficulties for the ERA writers in that the number of the chemical measurements could not be easily matched with toxicity and/or community structure information. I agree with Dr. Forbes that the “difference in sediment concentration trends (stations 4 – 8) between the benthic community samples (sediment PCB concentration declines) and the toxicity station samples (sediment PCB concentration increases) is unfortunate and does not increase the clarity of interpretation.”

This is clearly evident by the amount of time that the ERA authors spend trying to grapple with that specific issue – and in particular using the “most synoptic” data point from lab studies vs. the use of an average concentration of the field measurements. This appears to have been a *post priori* consideration – the Work Plan Appendix A.13 clearly defines analytical sampling procedures and the expectation that the sub-cores collected at each of the sites would be used for characterizing the analytical portion for all the legs of the triad. The “most synoptic” discussion is very difficult to understand as constructed in the document. What is very clear, and very important, is that the measured tPCB in the laboratory bioassays *must* be coupled with the bioassay results, and that the mean tPCB measurement be associated with the benthic infaunal data.

The variability in the field tPCB measurements confounds the construction of dose-response relationships between community structure and tPCBs. Figure D.2-10 presents the results of the 12 individual replicates measurements for tPCBs. Of particular interest, and concern for this assessment endpoint, is that the replicate tPCB concentrations span orders of magnitude. While troubling from a toxicological characterization standpoint – it is a fascinating finding; particularly within the high sand stations (1 – 5). At the December 11th meeting, the Panel questioned the presence of microcapsules of PCB-NAPL in the sand as a potential factor in the variability observed in the upstream, high sand stations. EPA has apparently confirmed that PCB microcapsule-NAPLs exist with sand grains, but we were not provided with specific data on locations. A question worth pursuing discussion within the Peer Review is whether these measurements can be taken to represent toxicologically relevant (i.e., bioavailable) concentrations. Porewater measurements at each of these sites might have provided a better means of assessing what is available to benthic infaunal communities, but those data are likely not available.

A second factor that will be important to evaluate is whether tPCB concentrations at each of the individual stations represent statistically different concentrations. As noted above, and I believe the Panel all agreed, the data should be used in their entirety, and the central question to ask first is if indeed the stations represent statistically different

concentrations. How the assessment of risk throughout this study, and in particular for the infaunal community data is open to re-evaluation based upon the answer to this question. “Eyeballing” the the data, it appears to me that that there will be at most only two to three groups of stations that have statistically different PCB concentrations. At this point, GE’s argument in the formal comments to the ERA is compelling: when plots of tPCBs for each replicate are made against taxonomic richness or organism abundance, a dose-response trend is not discernable. The Final ERA must take this into consideration.

For the moment then, I take position that use of the median value is not appropriate for assessing community data, nor for assessing bioassay results. The ERA presents bioassay results, and subsequently estimates MATCs, using median bulk sediment concentrations (see the presentation in Table D.3-2). I haven’t seen a compelling argument that this is appropriate: bioassay results should be based upon the measurements taken at the end of the test – not on a median calculated over all of the bioassay data sets.

Bedded Sediment and *In Situ* Bioassays Each of these tests is discussed below.

The in-lab *Hyalella azteca* tests are very solid data – there is excellent concordance between the results seen at 28, 35, and 42 day survival, the number of young produced, and overall weight gains. I agree with Dr. LaPoint that this is perhaps the most important data set produced within the benthic studies. Dose response is clearly demonstrated for tPCBs vs. 42-day survival and reproduction. One thing that was never made clear in the ERA (or the Work Plan or the EVS report) was why there are three separate time endpoints for this bioassay, and it would be helpful to make a clear explanation. As a corollary to that, in my opinion the only endpoint that should then be used for the MATC calculation is the 42-day survival and reproduction.

The in-lab *Chironomus tentans* is also a good test, although the fact that only the control and reference sediments showed any survival – and weight gains. This limits the utility of these data in calculating an MATC – although the argument is effectively made here that the NOEL for this test is 0.3 mg/kg tPCBs.

The conduct and results of the *in-situ* bioassays appears to be well-thought out and executed. The design and results are in general useful – although I do not believe that using *Daphnia magna* as a test organism for sediment-contact is appropriate. That is a water-borne exposure pathway; the Panel was generally in concurrence on this point.

Accepting that for the moment, the greater issue is presenting the results of the tests against “median” tPCB values over all the *in-situ* exposures. For example, Table 16 of the EVS presents the results for tPCBs for the 48 hour, 7 day, and 10 day exposures. In Table 17, the results of 48-hour sediment exposures are given. Using the results from Table 16, those data do not show a dose-response for *H. azteca* survival. In fact, at the highest concentration measured in sample 031 (522 mg/kg tPCBs), no response was

observed even though a response was observed at lower tPCB concentrations. Yet, in Table D.3-2, and in calculation of the MATC, a median bulk tPCB concentration of 77 mg/kg is presented. This looks to me like a “mix and match” of data that are convenient to a finding of risk, and do not follow accepted scientific methods. This becomes especially acute (no pun intended) in calculation of the final MATC.

The “mix and match” continues with the presentation of responses to the 10 day *in situ* exposures. Table 16 shows the 10-day tPCB sediment concentration for Station 428 to be 1.4 mg/kg, with an approximate 50% mortality response for *H. azteca*. No effect is shown at Station 019 at 14 mg/kg. Table D.3-2, however, shows a median tPCB concentration of 7.3 mg/kg tPCBs.

This continues over other data points and in my opinion is not appropriate. Results must be paired with the tPCBs that are measured for that test. Otherwise, it appears as if these data are skewed in a way that would point toward risk, instead of letting the data tell us if there is in fact risk present. This absolutely has to be the case for calculation of the MATC.

I want to make it very clear that I believe there is evidence within these data for PCB impacts on benthic infauna – but not within the way it is currently presented.

Benthic Infauna Community Structure Based on the Panel discussion, I am in agreement that other indices may be more appropriate, and should be explored. Dr. LaPoint indicated that Shannon-Weiner is not an appropriate index for systems where the ecology is dominated by few species, and recommend Simpson's dominance index. I do not believe that SW should be discarded, but that Simpson's may be more appropriate given the low number of taxa and the confounding course vs. fine-grained substrate. Dr. Oris noted that based on the SW index, there is a distinct response difference between fine grained and course grained sediments. Fine-grained species appear less sensitive than course grained species. I have reviewed that and concur; this needs to be explored further and discussed.

The ERA should do a better job of explaining why in the Work Plan soft depositional sediments were targeted for collection, but in the end six non-depositional (i.e., sand) sites were selected for sampling. Sampling of benthos differed somewhat for upstream coarse grained (shallow water) versus downstream fine-grained (from boat with fauna collected along shore therefore larger spatial separation in latter 10-20 m). This should be accounted for in the discussion of potential uncertainty. Otherwise, I note that the collection of infauna methods, enumeration, and calculation of indices all appear to be appropriately calculated and represented.

The overarching issue with the benthic data is the variability of tPCBs relative to the community analysis. As noted above, I suspect that when statistical analyses are applied to the within station tPCB variability, we may not see very much difference in these

stations. This is not accounted for in the ERA, or in the subsequent re-analyses presented by EPA in response to question TL12 (EPA Response to Peer Review Questions dated 12/11/03). I remain in agreement with two points in the evaluation of benthic community structure prepared by Scott Cooper of UCSB: (1) variability in tPCBs must be accounted for in this evaluation and (2) habitat factors need to be better accounted for in the benthic analysis.

Finally, there is no synthesis in the ERA between the bioassay data and the benthic infauna community data. The ERA determines an overall high toxicity for stations 7 and 8 based upon lab and *in situ* bioassay results, but also shows no community effects at the same stations. This contradiction needs more exploration and explanation than is currently given.

Tissue Analyses As noted above, benthic tissue analyses were originally intended for contributing data to the AQUATOX model, and appear to have been brought into the ERA *post priori*. Tissue concentrations can be an appropriate measurement endpoint – but it would be useful to have a discussion of what species were in the D-net collections that comprise “predators and shredders”. If I read the RFI correctly, historically, stoneflies and caddisflies have constituted these groups collected by similar methods. From an exposure standpoint – this is important.

MATC Calculation – The MATCs may not appropriately calculated for the following reasons: (1) the response data must be paired with tPCBs measured for that test, (2) only dose/response tests should be included, (3) only one test endpoint should be incorporated into the MATC (i.e., not include 28, 35 and 42 day endpoints), and (4) the MATC should be based upon chronic endpoints – not acute.

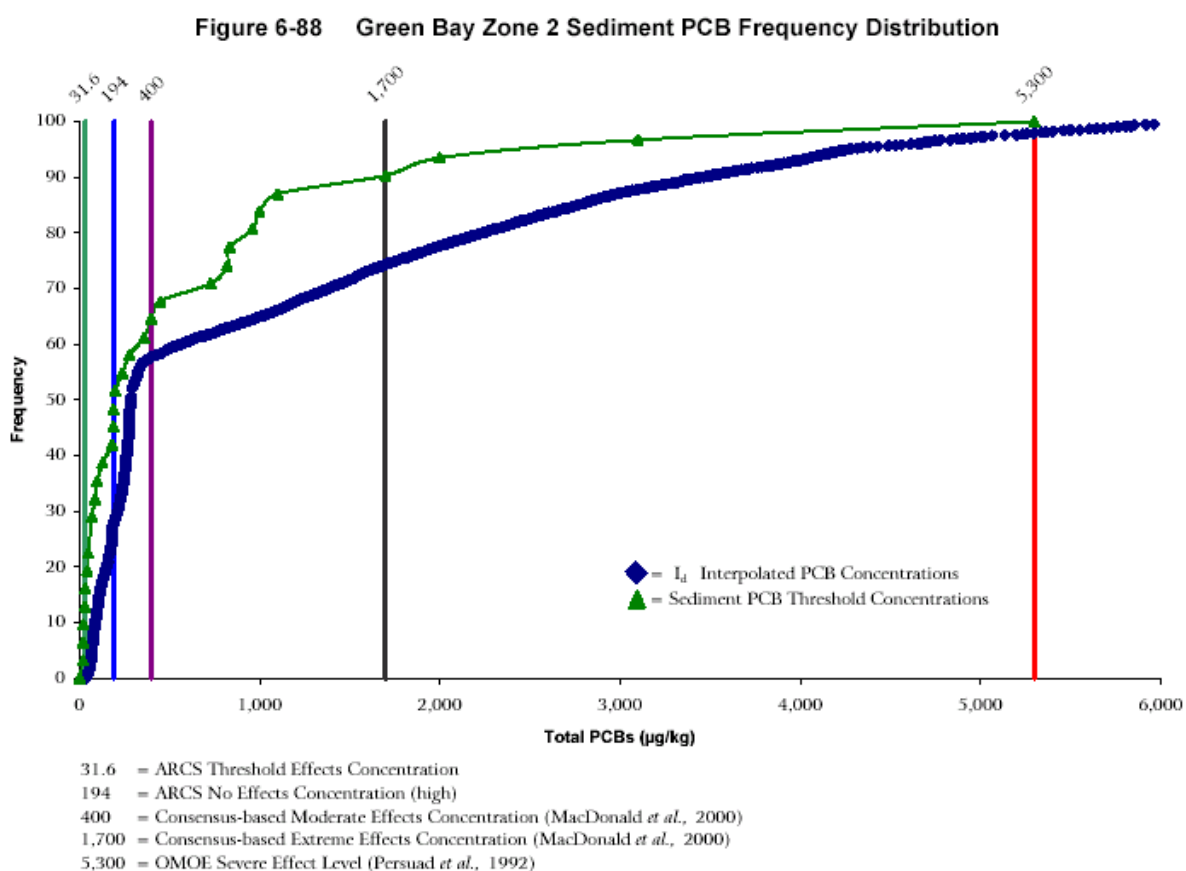
“(b) Were the GE studies and analyses performed outside of the framework of the ERA and EPA review (e.g., field studies) appropriate under the evaluation criteria, based on accepted scientific practices, and incorporated appropriately in the ERA?”

No GE studies were presented to us for consideration for this endpoint. However, the GE reanalysis, and the appendix presented by Dr. Scott Cooper, are appropriate and germane for consideration in the final ERA.

“(c) Were the estimates of exposure appropriate under the evaluation criteria, and was the refinement of analyses for the contaminants of concern (COCs) for each assessment appropriate?”

See comments under (a), above.

The ERA in made good use of probability distributions – and in particular the distribution of COC concentrations and the probability of exceeding the HQ (or MATC). This was not done in the case of benthic infauna, and should be conducted for each of the COC within the PSA. I presented an example that was used on the Lower Fox River and Green Bay PCB ERA, which is included below. These are simple frequency distributions of both measured, and GIS-base map interpolated tPCB concentrations. Something equivalent constructed for the PSA (and for the Connecticut reaches below the PSA) would provide a good visual presentation that is consistent with that done for other receptors in this ERA.



"(d) Were the effects metrics that were identified and used appropriate under the evaluation criteria?"

See comments under (a), above.

"(e) Were the statistical techniques used clearly described, appropriate, and properly

applied for the objectives of the analysis?"

See comments under (a), above.

The statistical methods seem generally appropriate. I concur with my colleagues who noted that the concerns raised by GE in response to the reanalysis of the benthic data are important. As noted by Dr. Forbes, "If a small fraction of the total variability in benthic species abundance can be explained by PCB concentration, despite statistical significance of the regression, this suggests that the role of PCBs in determining benthic community structure may be less important than concluded by EPA." This relates to the re-analysis done by Dr. Scott Cooper, and is relevant.

ERA could benefit from a better description of the statistical methods used. Enough detail should be presented so that the analyses could be repeated. A general recommendation: is that both effect size and significance are important and should be presented for all experimental results where appropriate. This is true throughout the ERA.

Shannon-Wiener may not be best measure of diversity for the sediments in which a few species dominate (Dr. Lapoint suggested Simpson's index).

"(f) Was the characterization of risk supported by the available information, and was the characterization appropriate under the evaluation criteria?"

Based upon the presentation in the draft ERA, the conclusions are not supported. To re-summarize, (1) the tests were not designed to assess overall benthic infaunal community health in the PSA – rather they were designed to identify dose/response relationships, (2) dose-response relationships were not evident in a "large number of endpoints" – as is claimed in the ERA (3) variability in the tPCB data need to be explored to determine if indeed the microscale differences observed are statistically different between the test stations, (4) the application of median calculations to the benthic population data or the bioassay data are not appropriate, (5) tissue residue data may not necessarily reflect sediment exposure, and that (6) the MATC is not appropriately estimated, which effects the risk estimates for the rest of the Housatonic River.

There was general concern noted by the Panel I concur with regarding the general risk terminology used for HQs; particularly a lack of definitions of low, moderate and high risk. HQs should be used as rough estimates of relative risk within assessment endpoints. Probability, or frequency of exceedance distributions, with defined boundaries for "low", "moderate" and "high" would be helpful in the final document. This needs to be expressed not only for tPCBs, but also for the other COCs.

"(g) Were the significant uncertainties in the analysis of the assessment endpoints identified and adequately addressed? If not, summarize what improvements could be made."

Principal uncertainty is within the tPCB sediment measurements, as is described above. Statistical significance and effect size are not considered here, or generally throughout the ERA. Power analyses are especially important, and I recommend that this be addressed. In addition, the ERA should provide a discussion of interlab and intralab variability on tPCB measurements, and the effect on uncertainty. At the least, the ERA should list which (how many) different labs provided tPCB measurements, methods for each lab, and provide an accounting for overall tPCB measurements per the discussion in Appendix D of the RFI.

"(h) Was the weight of evidence analysis appropriate under the evaluation criteria? If not, how could it be improved?"

No. See comments above.

"(i) Were the risk estimates objectively and appropriately derived for reaches of the river where site-specific studies were not conducted?"

The methodology for estimating risk in the lower Reaches is not inappropriate – I would submit that incorrect derivation of the MATC requires a reassessment of risks to the lower Housatonic. In particular, combining acute and chronic data to form a single MATC is not appropriate; these should be separate, and that was a general Panel consensus.

The Panel developed a series of “rules” that should be applied to the calculation of all MATCs for all Assessment Endpoints in this ERA. These include:

- Estimate separate acute and chronic MATCs,
- Multiple endpoints for the same receptor are not appropriate (e.g., dry-weight vs. ash-free dry weight)
- Only use the lab-measured “most synoptic data” for the bioassays
- Only those tests that display dose-response should be used in the MATC
- Use only those receptors relevant to the sediment (i.e., do not include Daphnia)
- Use all available data – not just the 6 lowest values
- If MATC is equal to or lower than the background/reference concentration, then use the background number.

"(j) In the Panel members' opinion, based upon the information provided in the ERA, does the evaluation support the conclusions regarding risk to local populations of ecological receptors?"

No, a reanalysis of the data is required before final conclusions can be made of the infaunal data. These include the spatial variability of tPCBs at each station, pairing the appropriate tPCB measurement with measurement endpoint, evaluation of substrate type on benthic infaunal communities, appropriate indices for infaunal communities, dominant taxa in the community structure, frequency distributions of tPCB and other COCs relative to TRVs and/or MATCs, and a better explanation of professional judgement used to assigning the weight-of-evidence factors.

Charge Question #5

“Are the summary discussions and conclusions in the ERA supported by the information provided in the report, and did the conclusions describe the risks in an objective, reasonable, and appropriate manner?”

See previous comments.

At this point, the data as presented do not support an MATC that could be used for a risk management decision. However, there are sufficient data, if properly re-evaluated along the lines the Panel recommended, that could form the basis for a supportable MATC.

Charge Question #6

“To the best of the Panel’s knowledge, is there other pertinent information available that was not considered in the ERA? Is so, identify the studies or data that could have been considered, the relevance of such studies or data, and how they could have been used in the ERA.”

There are specific gaps between the ERA Work Plan and the ERA that should be discussed in the ERA. These include:

- Dragonfly surveys – the Supplement Field Investigation Work Plan (“Work Plan) provides in Appendix A.7 a study to identify T&E dragonfly species within the Housatonic River. These species were identified in the PSA, but nothing else was done with that information. The ERA should explain why risks to these species was not identified as being important to the overall ERA.
- Freshwater mussels were identified as part of the FSP for the Housatonic River. Appendix A.8 stated that mussels “could potentially be directly or indirectly influenced by PCBs through several mechanisms, including: 1) direct mortality due to accumulation of PCBs in tissues, 2) impaired physiological function, 3) loss of food supplies, 4) loss of fish or amphibian intermediate hosts, and 5) loss or degradation of habitat.” Appendix A.15 is designed to identify toxicity to freshwater mussels. While we understand

from the presentations on December 18, 2003 to the Panel that these studies did not provide useable data, the ERA should document that finding in the interest of overall transparency.

- Crayfish studies were also listed under the Work Plan, yet no data were ever presented and discussed. Appendix A.16 specifically called for the co-location of crayfish samples with the benthic invertebrate samples. There are likely good reasons why these data were not collected or included – but they need to be discussed.

Finally, if there are additional sediment data that have been collected from behind the dams in Connecticut and not currently used in the risk estimates for the ERA, those need to be added to the ERA. The ERA may not conclude that there is no risk to the areas below the PSA, and the ERA should state that the study presentation cannot be a basis for a risk-management decision. The ERA can only qualitatively discuss relative levels of risk outside of the intensively-studies PSA.

Assessment Endpoint 2 Survival, growth and reproduction of amphibians

Charge Question #1

“Was the ecosystem of the Housatonic River watershed properly characterized, and was this information appropriately applied in the Problem Formulation and subsequently in the ERA?”

I am generally impressed with the comprehensive set of ecological and toxicological data assembled for this study. I have never seen a system better characterized for frogs in an eco risk assessment anywhere! This is an impressive collection of habitat and population data. The data was presented in the problem formulation in a logic and clear manner. Presentation of a life-history summary with Section 3, and in Appendix E, is something that should be done for all of the assessment endpoints. Studies undertaken by the EPA, and GE, clearly were designed to answer a long-overlooked, and data-poor area of toxicological research.

A consistent theme that I have raised throughout my comments is the very ambiguous formulation of the assessment endpoint, and “survival, growth and reproduction of amphibians was no exception. While not intended as an assessment endpoint, the “effects” box in Figure 4.1-1 was closer to what was probably intended: “Survival, growth, reproduction, and possible community alteration”. There was sufficient specificity within the measurement endpoint text box on page E-5 to understand what the intended study was undertaking.

As noted, the ecosystem characterization and Problem Formulation would benefit by clearer links between the two documents. A discussion of the process of selecting amphibians as a Receptor of Concern and the assessment endpoint formulation would greatly aid in the transparency of the overall document.

Charge Question #2

“Was the screening of contaminants of potential concern (COPCs), selection of assessment and measurement endpoints, and the study designs for these endpoints appropriate under the evaluation criteria?”

Given the general lack of amphibian tissue toxicity reference values, use of SQVs and AWQC are not an inappropriate way to screen for COPCs. However, since tissue measurements of Appendix 9 metals, PAHs, organo-chlorine pesticides were apparently made for field-collected frogs (based on a submitted chain-of-custody report), better use of these data should have been made. In general, the results shown in Appendix C to the *Rana pipiens* study showed non-detects for most organics, but DDE was detected, but not discussed.

Charge Question #3

"For each of the 8 assessment endpoints evaluated in the ERA (listed in Attachment B, and for which a specific Section and Appendix was prepared), address the following questions (discuss and label responses as 3.(assessment endpoint number).(question letter) for consistency):"

"(a) Were the EPA studies and analyses performed (e.g., field studies, site-specific toxicity studies, comparison of exposure and effects) appropriate under the evaluation criteria, and based on accepted scientific practices?"

A primary initial issue to address is the assessment and measurement endpoints – as the conduct of the studies relate directly to this. A general issue relates to use of maturation and development as an assessment endpoint. The Panel wanted to understand why “maturation” and “development” were important (as opposed to other indicators such as survival, sustainable populations, etc), and when queried, EPA responded that it was what the parties¹ agreed to. Several Panel colleagues were eloquent and passionate about “maturation” not being an assessment endpoint – arguing that if protect of populations of amphibians is the goal, then the assessment and measurement endpoints need to be appropriate parameters relating to the population (or societal values). Impairments can occur related to maturation, but that may not effect populations. I agree with their general comment, but believe that if the biological link is made between maturation and population – then that condition is effectively addressed. Recommend that the final ERA work to make that connection.

In general, the studies were well thought out, and executed. Unfortunately, not all of the studies proceeded as planned and I felt the study authors made the best use of data possible, even while not always agreeing with the conclusions. For example, a specific weakness cited in the Leopard frog reproductive success cited by GE and some of my Panel colleagues was that no frogs were captured from the reference area and that the study had to rely on purchased frogs for the control group. The suggestion has been made that since the reference group is not a true control it should be dropped from the statistical comparisons. I disagree with this position. I do not agree with this position; there is meaningful information to be derived from using the control frogs from outside the PSA. It adds a higher degree of uncertainty within the final conclusions, but recommend that those be retained in subsequent statistical analyses and assessment of risk.

"(b) Were the GE studies and analyses performed outside of the framework of the

¹ Although a formal BTAG was apparently not convened, I took “the parties” to mean the other federal (NOAA, USFWS) and State (Massachusetts and Connecticut) agencies. Recommend that a memo documenting the development of the Assessment and Measurement endpoints with the appropriate resource agencies and trustees be added as a memo or appendix to the Final ERA.

ERA and EPA review (e.g., field studies) appropriate under the evaluation criteria, based on accepted scientific practices, and incorporated appropriately in the ERA?"

GE's studies are well thought out and executed. The unfortunate aspect of their work is that it appears to have been conducted outside of the overall ERA Problem Formulation process, and more as a response (counterpoint) to what the ERA was testing. For example, the Leopard Frog study appears to have been appropriately designed to simply answer the question "are leopard frogs in the PSA and are egg masses present" -- as opposed to statistically powerful tests of whether leopard frogs are actively reproducing within the PSA. Notwithstanding the very small sample size, and the accompanying low statistical power, the study did demonstrate the Leopard Frogs are indeed in the PSA with associated egg masses. That is an effective counterbalance to the EPA study.

"(c) Were the estimates of exposure appropriate under the evaluation criteria, and was the refinement of analyses for the contaminants of concern (COCs) for each assessment appropriate?"

Generally, yes. Within the context cited under (a) and (b), above, the estimates of exposure are appropriate within both EPA and GE studies.

"(d) Were the effects metrics that were identified and used appropriate under the evaluation criteria?"

No further comments here.

"(e) Were the statistical techniques used clearly described, appropriate, and properly applied for the objectives of the analysis?"

Same general comment developed under the benthic infaunal section appropriate here -- recommend a better description of statistical techniques used in the ERA.

"(f) Was the characterization of risk supported by the available information, and was the characterization appropriate under the evaluation criteria?"

Within the context stated above, the general statements of risk from the field/lab studies appropriately support the risk statements made.

This is the first ROC where exposure modeling was incorporated, and I support the use of the models in this ERA. However, for this assessment endpoint the modeling leads to the extinction of amphibians in the PSA -- which appears to stretch credibility in that the two rarer species in the PSA are "present and accounted for". One recommendation is that

there be a better discussion of the parameterization of the models, and conducting a sensitivity analysis. For example, what set of parameter conditions would need to exist to not lead to extinction? Given the data collected, there should a discussion of the disconnect between model predictions and field observations

"(g) Were the significant uncertainties in the analysis of the assessment endpoints identified and adequately addressed? If not, summarize what improvements could be made."

As noted previously, a global issue is that for tissue analyses the intra-lab variability discussed in Appendix C.11-1 was 28.6%. The potential effect on the interpretation of risk should be included in the final ERA.

"(h) Was the weight of evidence analysis appropriate under the evaluation criteria? If not, how could it be improved?"

The same general comments from the benthos comments concerning the weighting of low, moderate and high risk apply to the amphibians as well. Otherwise, the final assessment of risk to amphibians appears to be appropriate – however with a high degree of uncertainty.

"(i) Were the risk estimates objectively and appropriately derived for reaches of the river where site-specific studies were not conducted?"

I concur with Panel colleagues who pointed out that the absence of samples from the vernal pools downstream of the PSA limits the ability to make statements about relative risks to amphibians.

Also, the MATC criteria developed for benthos are equally applicable to amphibians.

"(j) In the Panel members' opinion, based upon the information provided in the ERA, does the evaluation support the conclusions regarding risk to local populations of ecological receptors?"

In general, within a high degree of uncertainty, yes. However, I do not believe these data contain sufficient strength to formulate a supportable MATC around which a risk management decision can be constructed.

Charge Question #5

"Are the summary discussions and conclusions in the ERA supported by the information provided in the report, and did the conclusions describe the risks in an objective, reasonable, and appropriate manner?"

See answer to 4 (j) above.

Charge Question #6

“To the best of the Panel’s knowledge, is there other pertinent information available that was not considered in the ERA? Is so, identify the studies or data that could have been considered, the relevance of such studies or data, and how they could have been used in the ERA.”

Frequency of TRV exceedances could be estimated for amphibians, and would be helpful in presenting an assessment of risks within the PSA.

Assessment Endpoint 3. Survival, growth and reproduction of fish

Charge Question #1

“Was the ecosystem of the Housatonic River watershed properly characterized, and was this information appropriately applied in the Problem Formulation and subsequently in the ERA?”

For fish the ecosystem has been well characterized and, in general, appropriately applied to the Problem Formulation and subsequent ERA. As with all of the endpoints in this study, there is a tremendous volume of information that is presented in multiple documents. For example, ecological information on the largemouth bass may be found within the Work Plan, the RFI, the species profile for largemouth bass in the accompanying CD, in the Fish Biomass Estimate for the PSA, the GE study, and in Appendix F.2. It would greatly benefit the overall process to append or include the Species Profile – being sure to include all of the information from the various sources within Appendix F. The species profile should strive to include all of the GE habitat information (which it currently does not appear to), and be sure that life history information in Appendix F.2 is consistent with the species profile (see diet information).

A figure showing fish collection and tPCB concentrations within the PSA would be helpful

Charge Question #2

“Was the screening of contaminants of potential concern (COPCs), selection of assessment and measurement endpoints, and the study designs for these endpoints appropriate under the evaluation criteria?”

Screening of COPCs in the pre-ERA are appropriate under the evaluation criteria and are consistent with existing scientific knowledge and practice for ERAs.

What was not appropriately applied was pointed out under my general comments at the beginning of my review; assessment endpoints are ambiguously worded, and are not specific, testable hypotheses. As listed in EA Table 2.8.1, the assessment endpoint for fish is “survival, growth and reproduction”. In this form, the assessment endpoint allows for labeling any effect observed on these endpoints as being “adverse”. The risk question posed in Section 7 of the Work Plan offered somewhat more specificity;

Fish Community—Are ambient levels of PCBs present at concentrations in fish habitat sufficient to cause adverse effects in fish in the Housatonic River?

Even in that form, there is no means of supporting an “appropriate managerial decision”. As will be discussed below, the ERA does not make the distinction between potential individual effects as indicated in the laboratory toxicity studies, and a determination that there are ecologically significant effects to populations of fish in the PSA. As is aptly

done in the bird and wildlife sections, a statement of probability of exposure is coupled with a probability of effect. For example, for risk management purposes a statement such as “there is a 20% encountering PCB concentrations at which 50% of the exposed fry will experience mortality”, coupled with an understanding of field-measured populations, allows for setting what will be appropriate management actions for this endpoint. While probability of exposure is indeed presented, the MATC used is a mixture of lethal and subchronic effects in a fashion that cannot be used to integrate the data with population-level measurements.

There was considerable discussion among the Panel about the appropriateness of using the presence of disease and deformities as indicators of population health. On one level, the incidence of disease and deformities does not necessarily equate with population decline, but the presence does indicate presence of a stressor. I believe these data may be appropriately used, but either the assessment endpoints need to be broadened to include a “no mutated, diseased or ugly fish” rule on the River, or the measurement endpoint of reduced survival must be tightly linked to disease and deformities.

Charge Question #3

"For each of the 8 assessment endpoints evaluated in the ERA (listed in Attachment B, and for which a specific Section and Appendix was prepared), address the following questions (discuss and label responses as 3.(assessment endpoint number).(question letter) for consistency):"

"(a) Were the EPA studies and analyses performed (e.g., field studies, site-specific toxicity studies, comparison of exposure and effects) appropriate under the evaluation criteria, and based on accepted scientific practices?"

The EPA studies are solid investigations that are consistent with, or are on the cutting edge, of accepted scientific practice. This includes the fish egg exposure/toxicity studies conducted by Tillitt *et al* (2003), the fish biomass estimates (Woodlot 2002), and the numerous studies on fish tissue PCB levels that have been performed for the RCRA Facilities Investigation.

Specific issues that I have with these studies relate to (1) presentation of lipid-normalized data, (2) quality assurance/quality control issues relating to PCB and lipid measurements, (3) interpretation of exposure and effects, and (4) data omission/inclusion for the MATC calculation.

The presentation within the ERA makes the implicit assumption that PCBs are directly tied to lipid content without making that demonstration. While the literature argument for PCB “normalization” is made within the text, that confirmatory analysis is not made and should be considered. Distinct differences have been noted for different species and collection season between PCBs and lipid concentration in fish. In analysis of time trends for the Lower Fox River and Green Bay, a given species and sample type differed

between the reaches in one or more ways: 1) average PCB concentration differed, 2) time trend in PCB concentration differed, or 3) the relationship of PCB concentration to lipid content differed (TMWL and RETEC 2001). PCB associations with lipids were found to vary according to species, season collected, and area within the River or Bay PCB concentrations varied with percent lipid in different ways within two River reaches for two species of fish. Lumping across species does not take into account differences that have or have not occurred due to the well-known phenomenon of material transfer of lipid and contaminants between females and their eggs. For the Housatonic, the normalization may be appropriate – however it has not been demonstrated and needs to be done.

The occurrence of low lipid values (< 0.3%) in fish tissue samples is discussed in Appendix D to the RFI, but are not accounted for in the ERA. According to that appendix, this is especially true for fish fillet tissue samples, and that for some species these account for a large percentage of the existing data. This appears to be especially true for largemouth bass, yellow perch, and brown bullhead within Reaches 5 and 6. These data are used in the ERA, and a formula for deriving the whole body concentration from fillet data are presented in Section F.2.2, but lipid data are not discussed. As correctly noted in the RFI,

“When dividing similar fillet PCB concentrations by their corresponding lipid contents, the lipid values that are low by a factor of 10 will result in normalized PCB concentrations that are skewed high relative to the remainder of the data”

In light of the importance lipid normalization is given in the ERA, this uncertainty should be addressed and discussed.

The results of the fish studies are not presented in an objective fashion. For example, the text on Page F-31 discusses a range of effects that were observed in largemouth bass offspring in the PSA (“which were statistically different from those observed in the offspring from the reference location). These differences included reduced survival from hatch to swim up, development delays (increased days to swim-up), growth, and deformities. What is not covered is a discussion of how those related to tPCB concentrations reported in the whole adult largemouth bass in the ovaries.

That tPCB effects on fish are observed is not in dispute; rather it is the “selective” presentation of information that is called into questions. The table below presents three of those parameters versus tPCB concentrations from the Tillitt Phase 1 study. In the case of survival, there is clearly NO dose/response relationship; this is not discussed in the text. Rather, the ERA presents the negative effects observed without placing into context the observations. GE appropriately notes this within their comments, and I concur; as presented dose/response is not demonstrated. This does not meet standards for presentation of scientific information, and should be corrected.

tPCB Concentrations and Hatchout Parameters for LMB						
Tillitt Sample ID	Reach	Concentration tPCBs in Ovaries (Table 25 Mean ng/g ww)	Concentration tPCBs in Whole Bass (Table 24 Mean ng/g ww)	Survival to Swim-Up (Table 11 Mean)	Days from Hatch to Swim Up (Table 12 Mean)	Survival 15 days Post Swim-up (Table 15 Mean)
Three Mile Pond	Reference	634	106	68	7	28
Rising Pond	8	193,000	43,000	45	8	2
Woods Pond	6	251,000	108,000	57	9	26
Deep Reach	5C	315,000	149,000	28	8	13

This also continues over to the Phase 2 study (e.g., craniofacial abnormalities in injected largemouth bass).

"(b) Were the GE studies and analyses performed outside of the framework of the ERA and EPA review (e.g., field studies) appropriate under the evaluation criteria, based on accepted scientific practices, and incorporated appropriately in the ERA?"

The GE analysis is a very comprehensive picture of bass ecology and reproduction within the PSA. However, there study focused solely on demonstrating that bass spawning areas do exist within the PSA, and generally demonstrating that spawning occurs. That is very different from reproductive impairment, which is what EPA studies are intended to demonstrate. I concur with Dr. Oris the GE made no attempt to correlate the tPCBs in or near the nests. It's possible that since these areas were principally at the mouths of tributaries to the Housatonic, tPCBs would not accumulate there. A more convincing point would be to measure tPCBs within the sediments at the spawning areas. Having said that, egg "exposure" even in these areas will come from maternal transfer – not uptake from local sediment sources.

"(c) Were the estimates of exposure appropriate under the evaluation criteria, and was the refinement of analyses for the contaminants of concern (COCs) for each assessment appropriate?"

No. There is a discontinuity between the screening procedure applied in Appendix F, and additional measurements made under the Phase 1 Tillitt et al study that should be bridged. Section F.2.1 discusses the use of sediment and water quality values for screening out metals and pesticides as COC for fish. Given that pesticides in whole bass and ovaries were included in the Phase 2 study, those should be included in the screening assessment.

The statement within the ERA that PAHs in fish were not evaluated because these are metabolized is not an effective use of data or assessment of risk to fish. It is precisely those metabolic products that produce toxicity in fish species. Furthermore, it is inaccurate to state that PAHs are not measured in fish species. With sufficient exposure, carcinogenic PAHs have been measured in fish tissues. Better use of the PAH data should be made within the ERA.

"(d) Were the effects metrics that were identified and used appropriate under the

evaluation criteria?"

No. As noted above the ERA presents PCB “effects” without noting that those did not display dose/response. The absence of this discussion limits the utility of these data – and should be rectified. The mixing of endpoints in the MATC calculation, and the exclusion of other data, also limit the utility of this information.

"(e) Were the statistical techniques used clearly described, appropriate, and properly applied for the objectives of the analysis?"

No. See previous comments.

"(f) Was the characterization of risk supported by the available information, and was the characterization appropriate under the evaluation criteria?"

While I generally disagree with the use and assignment of toxicological and field study data in this assessment, the overall risk characterization is probably valid; While “there is evidence of harm for all measurement endpoints evaluated, but that the magnitude of those risks is generally low” is a qualitative statement, it is likely a valid assessment given the data.

Within the presentations of data for tPAH, I have to agree with the conclusions that risks attributable to PAHs appear low. However, I recommend this issue be re-evaluated.

Agree that the field surveys support an assessment that there are reproductive populations of large mouth bass within the Housatonic River – and believe that characterizing these results as “lack of catastrophic effects” is a negative disservice to the existing data. I do agree with the statement that the field surveys cannot predict “lesser impacts”, but believe those are not well defined in the ERA. The presence of disease and deformities should be put to better use – although it is my understanding that the incidence of these were relatively low. Finally, the population demographics is puzzling. The statement that the population is dominated by older individuals due to “lack of predation” appears speculative.

"(g) Were the significant uncertainties in the analysis of the assessment endpoints identified and adequately addressed? If not, summarize what improvements could be made."

The file conversion and associated lipid variability is not adequately characterized and should be improved in the final ERA. The issue of a complete representation of the data should also be expanded – particularly as it relates to MATC calculation and risk expression. The overall use of the field study data should be explored more thoroughly, and any uncertainties presented objectively.

"(h) Was the weight of evidence analysis appropriate under the evaluation criteria? If not, how could it be improved?"

See comments above.

Also, the GE field studies appear to be under-rated in the WOE. Recommend that either the weighting be adjusted upward for the field study, or a better discussion needs to be made on why those data are discounted.

"(i) Were the risk estimates objectively and appropriately derived for reaches of the river where site-specific studies were not conducted?"

No. While the methodology is sound overall, the estimation of the MATC should be revisited, based upon the same rules established for benthic infauna. Extrapolation to trout should be better explained – and perhaps consider application of uncertainty factors.

"(j) In the Panel members' opinion, based upon the information provided in the ERA, does the evaluation support the conclusions regarding risk to local populations of ecological receptors?"

As noted above, I generally agree that there are PCB-related risks to fish in the Housatonic River, but that the magnitude of those risks appear to be relatively low. I believe that the data exist within this study to actually represent both probability of exposure AND level of effect (e.g., 85% probability of encountering tPCB concentrations that yield a 20% swim bladder effect). The ERA should be re-expressed in those terms.

Charge Question #5

"Are the summary discussions and conclusions in the ERA supported by the information provided in the report, and did the conclusions describe the risks in an objective, reasonable, and appropriate manner?"

See previous comment

Charge Question #6

"To the best of the Panel's knowledge, is there other pertinent information available that was not considered in the ERA? Is so, identify the studies or data that could have been considered, the relevance of such studies or data, and how they could have been used in the ERA."

The absence of any reference or incorporation of the biological models – and in particular the diet parameterization is a big whole in the ERA. A considerable amount of work has

gone into developing food webs, bioaccumulation and transfer rates, age functions, growth dilution, lipid levels as a function of age, sex, and season, but none of that information is available or used in the ERA. Recommend that these two documents be reconciled before the release of the final ERA.

Assessment Endpoint 4. Survival, grown and reproduction of Insectivorous Birds

Charge Question #1

“Was the ecosystem of the Housatonic River watershed properly characterized, and was this information appropriately applied in the Problem Formulation and subsequently in the ERA?”

For insectivorous birds, the upland ecosystem appears in general to have been adequately characterized, and appropriately applied to the Problem Formulation and subsequent ERA. A deficit in the document is lack of transparency in providing links between the various studies conducted, and that the information collected and evaluated is not adequately represented in the ERA and Problem Formulation section.

Pertinent information is scattered through several documents, and is not drawn together into a single coherent summary– which in this case should be Appendix G. Documents that have to be accessed include the Work Plan (specifically Appendix A.24), Appendix G to the ERA, Final Report for tree swallows by Dr. Custer, the scope and final report on Robin Productivity prepared by ARCADIS for GE, and the landscape and species profiles prepared by Woodlot. There were also several instances where data were collected, and possibly used, but were not found anywhere in the ERA.

The rationale for placement of the nesting boxes in the Custer study is not covered in their summary document, but can be found in a careful read of Appendix A.24 in the Field Sampling Plan (Weston 2000). Appropriately, this included adequate tree swallow habitat and accessibility, but a very key consideration not made clear in the ERA reporting is that “the sites cover(ing) a range of sediment PCB contamination in the Lower River between the confluence and Woods Pond”. This is a good characteristic, but the selection process is not transparent and should be documented in the final ERA.

A corollary to the above is the lack of connectivity between other studies conducted under the ERA to this specific receptor group. For example, it appears that sediment PCB concentrations played a role in determining location of nest boxes, but this is not discussed anywhere in Appendix G or the supporting documents. Appendix A.20 to the Field Study Plan for the ERA (Weston 2000) states that “...WESTON, will collect sediment samples within the tree swallows’ expected foraging radius around the next boxes to support this study.” If such a study was conducted, it is not discussed in the ERA.

Tree swallow sightings listed in the Species profile are not discussed in Appendix G.

Aquatic insect collections made for the benthic invertebrate studies could/should be discussed in the context of prey items for tree swallows. For example, comparing Figure D.2-1 with Figure G.1-1 shows that benthic invertebrate sampling stations 1 – 3, 6, and 8

all occurred within nest-box locations. Benthic tissue measurements for tPCBs were made at each of these sites, and could/should be discussed within the context of tPCB measures made of prey items collected from juvenile birds.

There appears to be additional, important data collected regarding dietary intake for both species of birds evaluated that may not have been included or discussed in the ERA. For example, Slide 12 of Dr. Moore's presentation to the Peer Review Panel on December 18th for insectivorous birds listed dietary composition by dry mass for both tree swallow and American Robin. When questioned about these data, relative to the literature-based tables provided in Appendix G (Tables G.2-1 and G.2-26), Dr. Moore indicated that the data in his slide were site specific. Dr. Custer's independent report also indicates that "70% of dietary items of tree swallows from the Housatonic River were aquatic insects (unpubl. data)". If such data do exist, then the statement in the uncertainty section of Appendix G that states the "dietary compositions were derived from the reported in the literature for birds collected from other geographical locations" needs to be changed. If site specific data was indeed used, then I surmise that the data Dr. Moore used is perhaps that which Dr. Custer refers to in here report. When queried at the Panel meeting, EPA indicated that the unpublished insect data of Christine Custer is not in the ERA, but EPA has it. Recommend that these data be incorporated into the Final ERA.

The GE Robin study reports collecting nest locations, but those are not provided in a figure. How the resultant nests compare to soil PCB concentrations, and to tPCBs in the soil invertebrate collections made and shown in Figure G.1-2 should be made. A direct link between exposure and effects (or in this case, lack of effects) should be made.

This information is available, and the final ERA should draw links from the ERA to the supporting documents. For example;

- Discuss habitat preference and type within the Problem Formulation for both the American Robin and tree swallow. This may be more appropriate for Appendix G. Draw direct links from the Problem Formulation to the supporting information.
- In the text, link selection of nest-box study sites to the tree swallow species profile, and specifically Figure 2 in that profile.
- Provide specific information on how locations for nesting boxes in the tree swallow field studies were selected relative to the species profile (see more below on this)
- Provide a similar figure (Figure 2 – sightings) under the American Robin species profile.
- Provide a succinct rationale for selection of study for the American Robin, and provide a figure showing where the sampled nests were located, similar to Figure G.1-1
- Provide linkage to other assessment endpoints/studies conducted as part of the ERA. Specifically, sampling aquatic insects (including species and PCB

concentrations) and upland terrestrial invertebrates. Ensure that all data collected and used is fully disclosed and discussed in the ERA. (more below)

Charge Question #2

“Was the screening of contaminants of potential concern (COPCs), selection of assessment and measurement endpoints, and the study designs for these endpoints appropriate under the evaluation criteria?”

For insectivorous birds, the COPC screen, selection of assessment and measurement endpoints, and the resultant study design for these endpoints followed an objective and scientific process. The study plans and designs are in general within accepted scientific practice. Specific issues for insectivorous birds are discussed under Charge Question 3(a), below.

Charge Question #3

“For each of the 8 assessment endpoints evaluated in the ERA (listed in Attachment B, and for which a specific Section and Appendix was prepared), address the following questions (discuss and label responses as 3.(assessment endpoint number).(question letter) for consistency):”

“(a) Were the EPA studies and analyses performed (e.g., field studies, site-specific toxicity studies, comparison of exposure and effects) appropriate under the evaluation criteria, and based on accepted scientific practices?”

The lines of evidence provided by EPA were appropriate to this site, and are within the accepted practices of ecological risk assessments. Tree swallow field studies have been conducted at several PCB and other contaminated sediment sites; those studies are appropriately cited in the ERA. The mathematical models used to estimate intake and uncertainty are also well-established, and are accepted scientific practice within the risk assessment community. Field sampling procedures, quality assurance project plans, data quality objectives, and supporting study plans for ecological characterization and risk assessment were well articulated, and meet the requirements for data generation under both the RCRA, and the CERCLA authorities.

“(b) Were the GE studies and analyses performed outside of the framework of the ERA and EPA review (e.g., field studies) appropriate under the evaluation criteria, based on accepted scientific practices, and incorporated appropriately in the ERA?”

The line of evidence provided by GE for robin productivity within the PSA are appropriate to this site and are generally within the accepted practices of ecological risk assessment. However, the general lack of documentation concerning the study plan, quality assurance procedures, and DQOs limit the utility of the information provided. The overall study report could be improved by incorporating more specific information

from the ARCADIS Scope of Work to GE into the Final Report a discussion of how the areas to be studied were selected, a figure showing all nest sites, and those for which eggs and juveniles were collected for PCB analysis, would improve presentation.

The GE study collected very few eggs and nestlings for tPCB analyses, and does not articulate why so few samples were collected. This factor limits the strength of their argument that there is not a relationship between PCB concentration and nest success.

While the GE studies is one of the few studies included in the ERA which provided a formal statistical power analysis (provided in the ET& C 2003 publication), I would recommend EPA revisit those calculations. My own calculations on the data presented show a very low level of power – not at all consistent with what is presented in the ARCADIS Report.

In addition to the low power associated with few samples, there is also the unanswered question, why were only “viable eggs” sampled. A key element missing from this study (or at least is not documented) is why weren’t eggs randomly collected immediately after laying. Random collection would have eliminated what appears to be bias: testing only eggs that would hatch. I submit this limits the overall confidence with the robin study, and this uncertainty should be further explored in the final ERA.

A verification by EPA that the analytical portion of the study met the overall guidelines for quality assurance and control set forth in the QA/QC Plan (Weston 2000). The report also would be improved by a discussion of the uncertainties associated with the relatively few number of PCB analyses actually made.

An additional note is that the ARCADIS SOW to GE stated that in addition to tPCBs, “biota samples may be analyzed for PCB congeners, Appendix 9 pesticides...”. The ARCADIS scope is clear that it will depend upon “client preferences” whether these additional analyses were to be conducted. However, the final report does not make that distinction for us. If these analyses were done, then those data should be reported and discussed. If not, a simple note that the study focused solely on tPCBs would be helpful.

A general comment is that the overall ERA is done a disservice by not providing the ARCADIS report as an integrated part of the Appendices. The final ERA should include the GE report CDs as part of the ERA package.

“(c) Were the estimates of exposure appropriate under the evaluation criteria, and was the refinement of analyses for the contaminants of concern (COCs) for each assessment appropriate?”

No. There is a discontinuity between the screening procedure applied in Appendix G, and the field study reports that should be bridged. Section G.1 states that the other COPCs,

and particular the organochlorine pesticides, are screened out “because their actual concentrations in the PSA were likely much lower than the measured values due to laboratory interference”. However, the Custer report includes analyses and detection of polycyclic aromatic hydrocarbons (no detections) and aliphatic hydrocarbons, metals, and several organochlorine compounds including DDT and derivatives. That appendix presents a credible argument for why these other detected COPCs are not at levels of environmental concern in tree swallow tissues. The ERA should be amended to incorporate those findings as a reason for refining the COC list for the probabilistic risk assessment.

“(d) Were the effects metrics that were identified and used appropriate under the evaluation criteria?”

Yes. No further comment here.

“(e) Were the statistical techniques used clearly described, appropriate, and properly applied for the objectives of the analysis?”

Yes, with two exceptions noted below.

The ERA reports that there is a small, but statistically significant negative correlation between tPCBs and hatching success in tree swallows. This is noted throughout the document including the Executive Summary, is noted in the Quantitative Measure attribute, and the Risk Summary. However, it is somewhat misleading, and is somewhat incongruous inconsistent with the analysis presented by Dr. Custer in her report. To quote:

“There was a negative relationship between total PCBs and hatching success in 1998 and 1999 combined ($P = 0.049$, Fig.4), ***but not when 2000 was added to the model ($P=0.273$)*** [emphasis added].

The fact that 1998 and 1999 were significant when combined is very important – and better use can be made of the data. The statistical relationship when the 2000 data is added is tenuous at best, and I would not go as far as to even label it as a “weak” relationship.

While I agree with the ERA’s position that the statistical analysis conducted in the ARCADIS report were likely not the appropriate choice given the limited data size. However, if the ERA is going to find “fault” with the statistical tests employed, those need to be articulated. In my opinion, given that the sample sizes are small, and the test between the PSA and reference nests are unequal sample size, the more appropriate test to use would have been the Kruskal-Wallis test for non-parametric comparison, but I suspect given the very low difference between means, this too would show now difference.

As noted, I cannot repeat the optimistic power from the ARCADIS report data, and have recommended that EPA look at those data themselves. I would further submit that a power analysis of the tree swallow study would also be required for the reproductive metrics only.

"(f) Was the characterization of risk supported by the available information, and was the characterization appropriate under the evaluation criteria?"

Based on the evidence presented, the ERA in my opinion correctly assigns a “low risk” from exposure to tPCBs and TEQs to insectivorous birds. While having some general concerns with the assignment of “weights” to the metrics, the ERA correctly relied on the field study data to ultimately assess risk.

"(g) Were the significant uncertainties in the analysis of the assessment endpoints identified and adequately addressed? If not, summarize what improvements could be made."

An important source of uncertainty not discussed is the measurement error that likely occurred in PCB measurements between the two field studies. Data variability is a consistent issue throughout the entire ERA and RCRA Facilities Investigation. Measurement error is cited in Section 7.3 of the Work Plan (Weston 2000) as a source of uncertainty. This uncertainty is not specifically called out or accounted for in Appendix G.

An evaluation of data quality and interpretation issues is presented in Appendix D to the RFI. While that is not specific to bird tissues, the fact that the RPDs in fish tissue duplicates and splits is reported to be approximately 29% for intra-lab comparisons, and 41% for inter-lab comparisons would likely also apply to analyses of invertebrate (prey), egg, and nestling tissues. While this variability cannot be controlled at this stage of the ERA, accounting for it in the uncertainty section is important – recognizing that it will most likely be a qualitative discussion. Having said that, noting that EPA and GE have split samples for several media, I would be surprised if this was not done for these studies, as well. If that is the case, then lab variability must be accounted for explicitly.

"(h) Was the weight of evidence analysis appropriate under the evaluation criteria? If not, how could it be improved?"

For insectivorous birds, there are several instances where the weight assignments do not appear to meet the criteria of “objectivity, consistency, and reasonableness”.

For the analysis for *stressor/response*, the ERA assigns the EPA study a “moderate/high” value, but assigns the GE study only a “moderate” value. The principal reason, as worded in the text, is that the study did not “demonstrate a relationship between tPCB

exposure and reduced reproductive output”. The assignment, then, is not *objective*: lack of a finding is not cause for a reduced ranking. GE’s assertion that robins do not appear to be affected, within the test ranges measured, is supported by the ARCADIS study. The ERA secondarily sites the “nature of the study design ... is not suited to assessing stressor-response relationships”. That there are differences in the strengths of the two study in terms of control over where nest sites were located (we are not provided a figure on where the monitored nests are in the ARCADIS study – something that should be corrected in the next draft), the number of analyses on eggs and fledglings, the relative statistical power associated with the increased number of samples, is clear, and may provide better strength to the ranking. However, the ERA has already assigned the GE-study a “high value” for the Degree of Association.

It is my opinion that the modeled exposure and effects endpoint was inappropriately assigned a moderate weighting value. The strength of any model lies in the correlation of predicted and observed effects. Table G.4-7 and G.4-8 show that the modeled exposure predicted a high magnitude of risk for tree swallows, and an intermediate risk for robins. Both field studies demonstrated, and were appropriately assigned low magnitudes of risk. The conclusion is that the model over-predicts risk for these species, and those should be assigned a low weighting value. The moderate assignment may be a residual of an *a priori* assignment given within the Work Plan, which stated:

“Assuming that reasonable input parameters and distributions are applied to the modeling effort, a Medium weight-of-evidence will be given to this measurement endpoint.” (Weston 2000, Section 7, page 7-76, lines 7 – 8)

"(i) Were the risk estimates objectively and appropriately derived for reaches of the river where site-specific studies were not conducted?"

Derivation of the upper and lower effects thresholds for both Aroclors and TCDD-TEQs (Figures G.3-1 and .3-2), while conservative, are consistent with both scientific and ERA practice. Application of those thresholds to the lower reaches of the Housatonic, based on these conservative thresholds, show low risk and are appropriate.

Again, would recommend applying the MATC “rules” established for benthic infauna here as well. In the case of insectivorous birds, however, I do not believe that will significantly alter the results.

"(j) In the Panel members’ opinion, based upon the information provided in the ERA, does the evaluation support the conclusions regarding risk to local populations of ecological receptors?"

For insectivorous birds, the lines of evidence and data presented support the conclusion of the ERA.

Charge Question #5

“Are the summary discussions and conclusions in the ERA supported by the information provided in the report, and did the conclusions describe the risks in an objective, reasonable, and appropriate manner?”

Yes, within the context of specific issues related to statistical analysis, low power associated with the robin field study, and uncertainties, cited above.

Charge Question #6

“To the best of the Panel’s knowledge, is there other pertinent information available that was not considered in the ERA? Is so, identify the studies or data that could have been considered, the relevance of such studies or data, and how they could have been used in the ERA.”

No. The field studies and models applied within this study are complete and appropriate.

Assessment Endpoint 5. Survival, growth and reproduction of piscivorous birds

Charge Question #1

“Was the ecosystem of the Housatonic River watershed properly characterized, and was this information appropriately applied in the Problem Formulation and subsequently in the ERA?”

Yes, with noting that the same comments raised on the previous four assessment endpoints are applicable here, as well.

Charge Question #2

“Was the screening of contaminants of potential concern (COPCs), selection of assessment and measurement endpoints, and the study designs for these endpoints appropriate under the evaluation criteria?”

Screening of COPCs were appropriate for this receptor.

The same general comment concerning transparency of selection of this assessment endpoint – and the accompanying measurement endpoint, with the resource trustees is applicable here.

EPA’s study design using principally modeled exposures is appropriate and has been previously commented on. The documented presence of osprey within the PSA demonstrates the appropriateness of utilizing this species as a receptor of concern.

As has been noted previously, the GE field studies were not part of the *a priori* determination of measurement endpoints, and here to the principal aim was solely to document presence of reproducing kingfishers – and is not necessarily a demonstration that reproductive impairment is not occurring.

Charge Question #3

“For each of the 8 assessment endpoints evaluated in the ERA (listed in Attachment B, and for which a specific Section and Appendix was prepared), address the following questions (discuss and label responses as 3.(assessment endpoint number).(question letter) for consistency):”

“(a) Were the EPA studies and analyses performed (e.g., field studies, site-specific toxicity studies, comparison of exposure and effects) appropriate under the evaluation criteria, and based on accepted scientific practices?”

Yes – the modeled exposures were appropriate and are based on accepted scientific practices. Some reconciliation should be made on the conclusions of reproductive risks

to osprey, with the field observation that sitings within PSA have increased in numbers from 10, to 190.

"(b) Were the GE studies and analyses performed outside of the framework of the ERA and EPA review (e.g., field studies) appropriate under the evaluation criteria, based on accepted scientific practices, and incorporated appropriately in the ERA?"

I agree with my colleagues in noting that the overall study design and techniques were very good, and remarkable in their presentation. As noted above, the study at best demonstrates that reproducing kingfishers exist within the PSA, and contributes little beyond that. The few number of sampled nests, and the single season, contribute to very low power associated with the conclusions. Having said that, the kingfisher study is the only site-specific field data for this assessment endpoint, and is appropriately incorporated into the ERA.

"(c) Were the estimates of exposure appropriate under the evaluation criteria, and was the refinement of analyses for the contaminants of concern (COCs) for each assessment appropriate?"

In general, yes. I would like to have seen more use made of the crayfish data within the modeling approach – particularly since crayfish are directly exposed to sediments. Use of the extensive fish data is appropriate.

"(d) Were the effects metrics that were identified and used appropriate under the evaluation criteria?"

Yes.

"(e) Were the statistical techniques used clearly described, appropriate, and properly applied for the objectives of the analysis?"

Within the limits of previous comments made about needing better descriptions of statistical methods, yes.

"(f) Was the characterization of risk supported by the available information, and was the characterization appropriate under the evaluation criteria?"

Within the context of the limited information presented in the ERA (modeling and low power with the kingfisher study), the assessment of risk appears to be appropriate.

"(g) Were the significant uncertainties in the analysis of the assessment endpoints identified and adequately addressed? If not, summarize what improvements

could be made."

Uncertainties in the modeled projections were generally described. The low power associated with one season and 6 nests within the kingfisher study should be more clearly articulated in the final document.

"(h) Was the weight of evidence analysis appropriate under the evaluation criteria? If not, how could it be improved?"

For osprey, it may have been more appropriate to divide the PSA into north and south based upon the river characteristics. The narrow, relatively fast moving water is not going to be conducive to osprey hunting, which is more likely to occur within the ponded areas in the south. While this may seem inconsistent with the rest of the document, from an exposure standpoint it makes sense.

"(i) Were the risk estimates objectively and appropriately derived for reaches of the river where site-specific studies were not conducted?"

Within the limits of uncertainty associated with model projections, yes.

"(j) In the Panel members' opinion, based upon the information provided in the ERA, does the evaluation support the conclusions regarding risk to local populations of ecological receptors?"

Within the limits of the information, these are the only conclusions one could draw.

Charge Question #5

"Are the summary discussions and conclusions in the ERA supported by the information provided in the report, and did the conclusions describe the risks in an objective, reasonable, and appropriate manner?"

No comment here.

Charge Question #6

"To the best of the Panel's knowledge, is there other pertinent information available that was not considered in the ERA? Is so, identify the studies or data that could have been considered, the relevance of such studies or data, and how they could have been used in the ERA."

No further comments here.

Assessment Endpoint 6. Survival, growth and reproduction of piscivorous mammals

Charge Question #1

“Was the ecosystem of the Housatonic River watershed properly characterized, and was this information appropriately applied in the Problem Formulation and subsequently in the ERA?”

Yes. It would be appropriate in the Problem Formulation to discuss mink habitat and food preferences, and to show in the PF a figure with mink sightings. All the information exists, but it is scattered through several documents. A single summary would be helpful.

Charge Question #2

“Was the screening of contaminants of potential concern (COPCs), selection of assessment and measurement endpoints, and the study designs for these endpoints appropriate under the evaluation criteria?”

Yes – this is a good study with solid consideration of the study design and intended use

Charge Question #3

“For each of the 8 assessment endpoints evaluated in the ERA (listed in Attachment B, and for which a specific Section and Appendix was prepared), address the following questions (discuss and label responses as 3.(assessment endpoint number).(question letter) for consistency):”

“(a) Were the EPA studies and analyses performed (e.g., field studies, site-specific toxicity studies, comparison of exposure and effects) appropriate under the evaluation criteria, and based on accepted scientific practices?”

Yes, the combination of solid field studies, with the caged mink feeding exposure, is a solid piece of work that follows (frankly, exceeds!) general scientific/toxicological practice.

One issue to consider further is whether the model dataset was biased by only using carp and goldfish as the prey species. Mink are opportunistic, and will take a variety of prey species, but carp likely represent the highest levels of PCBs in fish in the River. A good discussion in the uncertainty section on this issue would probably address that concern.

Kudos for the congener work. That was excellent and helpful. Better use of the other organochlorine and metals data would also be recommended.

"(b) Were the GE studies and analyses performed outside of the framework of the ERA and EPA review (e.g., field studies) appropriate under the evaluation criteria, based on accepted scientific practices, and incorporated appropriately in the ERA?"

The same general comment raised previously concerning the GE studies are valid here. The work is solid and appears to have been lead by competent scientists. The study lacks a formal objective beyond demonstrating mink can be found within the PSA. This was clearly demonstrated, but not much else can be derived from their work. The issue of whether the GE field personnel were appropriate, and how gloves/rubber boots effect scent posts is well outside of my field of knowledge – I leave others to comment on this.

"(c) Were the estimates of exposure appropriate under the evaluation criteria, and was the refinement of analyses for the contaminants of concern (COCs) for each assessment appropriate?"

My only comment here is that the final ERA account for GE's comment that there is no dose/response relationship in the reproductive caged-mink endpoints (see GE Figure 53). If there is no statistical differences between stations, and the relationship is not strictly dose dependent, than it is difficult to make a case that PCBs are impacting in the PSA.

"(d) Were the effects metrics that were identified and used appropriate under the evaluation criteria?"

The issue of jaw lesions should be further explored. While I concur that connecting the presence of jaw lesions leading to starvation is speculative, there is some evidence cited in the scientific literature in the ERA. It's also important to determine if the observed effects are hyperplasiac, or precancerous – the difference is important. Having said that, I do believe that the presence of these lesions can be supportive of risks in the PSA to mink.

"(e) Were the statistical techniques used clearly described, appropriate, and properly applied for the objectives of the analysis?"

The statistical methods were not provided in the sections. Same global comment for all of the sections on statistical methods.

"(f) Was the characterization of risk supported by the available information, and was the characterization appropriate under the evaluation criteria?"

With the caveat of needing to look again at GE's comments on dose/response – this is a very well done study and supports the conclusions of the ERA.

"(g) Were the significant uncertainties in the analysis of the assessment endpoints identified and adequately addressed? If not, summarize what improvements could be made."

The uncertainty in this section was well characterized. Same global comments regarding analytical variability are applicable here.

"(h) Was the weight of evidence analysis appropriate under the evaluation criteria? If not, how could it be improved?"

See previous comments.

"(i) Were the risk estimates objectively and appropriately derived for reaches of the river where site-specific studies were not conducted?"

Same criteria from Assessment Question 1.

"(j) In the Panel members' opinion, based upon the information provided in the ERA, does the evaluation support the conclusions regarding risk to local populations of ecological receptors?"

Yes.

Charge Question #4

"Are the summary discussions and conclusions in the ERA supported by the information provided in the report, and did the conclusions describe the risks in an objective, reasonable, and appropriate manner?"

Statements such as "significant" and "catastrophic" should be eliminated. Likewise, "unacceptable" is a value judgement, and should not be in the ERA. State the levels of risk, or alternatively, define what is "unacceptable" upfront.

As drafted, the ERA needs to consider some revisions and resynthesis. I believe that with the reanalysis of concerns cited by the Peer Review, the ERA will be supportable.

Charge Question #5

"To the best of the Panel's knowledge, is there other pertinent information available that was not considered in the ERA? Is so, identify the studies or data that could have been considered, the relevance of such studies or data, and how they could have been used in the ERA."

No.

Assessment Endpoint 7. Survival, growth and reproduction of omnivorous and carnivorous mammals

Charge Question #1

“Was the ecosystem of the Housatonic River watershed properly characterized, and was this information appropriately applied in the Problem Formulation and subsequently in the ERA?”

Same comments from AEs 1 – 4 apply here. No further comment.

Charge Question #2

“Was the screening of contaminants of potential concern (COPCs), selection of assessment and measurement endpoints, and the study designs for these endpoints appropriate under the evaluation criteria?”

Same comments from AEs 1 – 4 apply here. No further comment.

Charge Question #3

“For each of the 8 assessment endpoints evaluated in the ERA (listed in Attachment B, and for which a specific Section and Appendix was prepared), address the following questions (discuss and label responses as 3.(assessment endpoint number).(question letter) for consistency):”

“(a) Were the EPA studies and analyses performed (e.g., field studies, site-specific toxicity studies, comparison of exposure and effects) appropriate under the evaluation criteria, and based on accepted scientific practices?”

Studies appear to be appropriate for the risk assessment and are based on accepted scientific practices.

“(b) Were the GE studies and analyses performed outside of the framework of the ERA and EPA review (e.g., field studies) appropriate under the evaluation criteria, based on accepted scientific practices, and incorporated appropriately in the ERA?”

This is not my area of expertise. Having said that, the study appears to be well done, and follows solid scientific practice and principle. The largest outstanding issue that should be addressed is the apparent discrepancy between GE’s analysis, and the re-analysis conducted by EPA. GE’s reanalysis of EPA’s reanalysis using same data and technique gave a non-significant result. GE/Dr. Boonstra treated the grids as replicates and did not weight them differentially, as a function of sample size within grids. They used probit transformation following the methods of Sokal and Rolf (pp 544-547). Grid 3 was not included in the analysis of males because there was only one male. EPA followed the

method of Baylor & Oris (1997; ET &C); general linear modeling with probit transformation. The grids were weighted according to total number of organisms in the treatment.

Summarizing from Dr. Samples notes (which I concur with), while the analyses are different, they are both likely valid. Both are probit regressions; simple probit (Boonstra rework) is a monotonic linear model, while the Baylor and Oris model can account for non-linear response and account for error bounds differently. Bottom line is that the results indicate that response is not strong – both methods should probably be reported and discussed in final ERA.

"(c) Were the estimates of exposure appropriate under the evaluation criteria, and was the refinement of analyses for the contaminants of concern (COCs) for each assessment appropriate?"

See comment above.

"(d) Were the effects metrics that were identified and used appropriate under the evaluation criteria?"

The literature summary is impressive and appears substantively complete. The only comment I had was that there did not appear to be any canine studies with PCBs; I found that surprising.

"(e) Were the statistical techniques used clearly described, appropriate, and properly applied for the objectives of the analysis?"

Previous comments regarding transparency relevant here.

"(f) Was the characterization of risk supported by the available information, and was the characterization appropriate under the evaluation criteria?"

Yes.

"(g) Were the significant uncertainties in the analysis of the assessment endpoints identified and adequately addressed? If not, summarize what improvements could be made."

Within the limits of the data, uncertainties fairly well characterized. Same general previous comments re. uncertainty apply here as well.

"(h) Was the weight of evidence analysis appropriate under the evaluation criteria? If not, how could it be improved?"

Yes – although I would recommend that a separate table be constructed for foxes and for shrews.

"(i) Were the risk estimates objectively and appropriately derived for reaches of the river where site-specific studies were not conducted?"

No estimates of risk were made outside of the PSA. If those are developed for the final ERA, the same general rules for MATC development apply here.

"(j) In the Panel members' opinion, based upon the information provided in the ERA, does the evaluation support the conclusions regarding risk to local populations of ecological receptors?"

Yes. Given the site-specific data, along with an abundance of literature data, sufficient information exists upon which to develop and MATC and formulate an appropriate risk management decision for this receptor group.

Charge Question #5

"Are the summary discussions and conclusions in the ERA supported by the information provided in the report, and did the conclusions describe the risks in an objective, reasonable, and appropriate manner?"

Yes.

Charge Question #6

"To the best of the Panel's knowledge, is there other pertinent information available that was not considered in the ERA? Is so, identify the studies or data that could have been considered, the relevance of such studies or data, and how they could have been used in the ERA."

I am aware of no other pertinent or relevant information that could be applied here.

Assessment Endpoint 8: T&E Specs

Charge Question #1

"Was the ecosystem of the Housatonic River watershed properly characterized, and was this information appropriately applied in the Problem Formulation and subsequently in the ERA?"

General comments given on the other assessment endpoints are applicable here. Would be helpful to have a discussion of eagle sitings and use of the River in the Problem Formulation for this species.

Charge Question #2

"Was the screening of contaminants of potential concern (COPCs), selection of assessment and measurement endpoints, and the study designs for these endpoints appropriate under the evaluation criteria?"

Yes.

Charge Question #3

"For each of the 8 assessment endpoints evaluated in the ERA (listed in Attachment B, and for which a specific Section and Appendix was prepared), address the following questions (discuss and label responses as 3.(assessment endpoint number).(question letter) for consistency):"

"(a) Were the EPA studies and analyses performed (e.g., field studies, site-specific toxicity studies, comparison of exposure and effects) appropriate under the evaluation criteria, and based on accepted scientific practices?"

Yes – modeling issues have been dealt with previously

"(b) Were the GE studies and analyses performed outside of the framework of the ERA and EPA review (e.g., field studies) appropriate under the evaluation criteria, based on accepted scientific practices, and incorporated appropriately in the ERA?"

No GE studies were performed for this receptor group.

"(c) Were the estimates of exposure appropriate under the evaluation criteria, and was the refinement of analyses for the contaminants of concern (COCs) for each assessment appropriate?"

Yes.

"(d) Were the effects metrics that were identified and used appropriate under the evaluation criteria?"

Within the limits of modeling discussed in previous section, the evaluation criteria are appropriate.

"(e) Were the statistical techniques used clearly described, appropriate, and properly applied for the objectives of the analysis?"

Within the limits of modeling discussed in previous section, the evaluation criteria are appropriate.

"(f) Was the characterization of risk supported by the available information, and was the characterization appropriate under the evaluation criteria?"

Yes, within the limits of uncertainty the assessment of risk is appropriate.

"(g) Were the significant uncertainties in the analysis of the assessment endpoints identified and adequately addressed? If not, summarize what improvements could be made."

The final ERA should address how the limits of analytical variability could affect the overall determination of risk.

"(h) Was the weight of evidence analysis appropriate under the evaluation criteria? If not, how could it be improved?"

Yes. However, I concur with my colleagues who noted that there should be different risk weighting for different life stages. Particularly in the fact that eggs were at high risk, adults at low risk, and the final assignment was "moderate". This is probablomatic in that eggs are a more sensitive life stage than adults.

"(i) Were the risk estimates objectively and appropriately derived for reaches of the river where site-specific studies were not conducted?"

Yes. My only comment is that the fact that suitable habitat exists for eagle nesting, but nests are not found within the PSA is not in and of itself an indicator of risk. This can, however, be used as a pillar supporting the overall assessment of risk.

"(j) In the Panel members' opinion, based upon the information provided in the ERA, does the evaluation support the conclusions regarding risk to local populations of ecological receptors?"

Within the limits of modeling uncertainty, yes.

Charge Question #5

“Are the summary discussions and conclusions in the ERA supported by the information provided in the report, and did the conclusions describe the risks in an objective, reasonable, and appropriate manner?”

Charge Question #6

“To the best of the Panel’s knowledge, is there other pertinent information available that was not considered in the ERA? Is so, identify the studies or data that could have been considered, the relevance of such studies or data, and how they could have been used in the ERA.”