Human Health Risk Assessment
GE/Housatonic Site
Rest-of-the-River

Peer Review Report
December 17, 2003

John C. Kissel, Ph.D.
Seattle, WA

Preface: Comments presented here reflect time-constrained evaluation of a multi-volume risk assessment, numerous background documents and multiple oral presentations. No assumptions should be inferred regarding approval or disapproval of any matter not explicitly addressed herein.

A. Phase 1 – Direct Contact Exposure Screening Were the procedures used in Phase 1 of the HHRA to screen out properties and areas from further evaluation as well as the application of those procedures appropriate under the evaluation criteria? In addressing this question, consider:

- the general procedures used;
- the SRBCs used for the COPCs; and
- the land use and exposure categories considered and the classification of particular parcels and areas into those categories.

EPA has developed guidance for soil screening at Superfund sites. However, I can find no mention in the Rest-of-River Human Health Risk Assessment (ROR HHRA) of EPA's Soil Screening Guidance (OSWER 9355.4-23, July 1996) or the more recent Supplemental Guidance For Developing Soil Screening Levels for Superfund Sites - Peer Review Draft (OSWER 9355.4-24, March 2001). Given that those documents were produced by OSWER specifically for this type of application, it is unclear why they are not cited. Potentially relevant guidance related to future land use (Land Use in the CERCLA Remedy Selection Process (OSWER Directive No. 9355.7-04, 1995) is also not cited.

Screening procedures should be conservative and generally err on the side of retention rather than exclusion of exposure areas. In three respects the screening procedure appears to be neither conservative nor consistent with guidance cited above. These are 1) Use of variable target risk levels (from 1 to $5 \cdot 10^{-6}$) rather than a single default value of 10^{-6}; 2) Consideration of direct contact only when screening agricultural lands when the SGDSSL appears to indicate that other pathways should be considered; and 3) consideration of current land uses only.

B. Phase 2 – Direct Contact Exposure Assessment

1. Were the following aspects of the direct-contact exposure assessment appropriate under the evaluation criteria?

- The exposure scenarios which were evaluated.
- The exposed populations which were selected for each scenario.
- The exposure areas identified based upon potential current and future use(s).
- The routes of exposure for each scenario.

Consider the following when addressing this question:

> Current and reasonably anticipated future land uses, physical conditions, and accessibility;
> Locations, concentrations, and distribution of COPCs in the sediment, bank soil, and floodplain soil; and
> Ages of the selected exposed populations.

The direct contact scenarios considered generally appear appropriate and adequately inclusive.

2. Have the most important exposure pathways been identified and evaluated?

In the context of direct contact exposure, the pathways considered generally appear appropriate and adequately inclusive.

3. Were the approaches and methods used to calculate and apply exposure point concentrations (EPCs) for the direct-contact exposure assessment appropriate under the evaluation criteria?

The methods used to calculate and apply soil EPCs are not entirely transparent although, to the extent that they can be understood, they do seem consistent with EPA guidance. The use of weighting factors related to exposure point accessibility seems arbitrary. The assumption that exposure is random within exposure sub-areas is also questionable. An individual could habitually contact soils that are more or less contaminated than the mean as a result of selective visitation within a parcel.

EPA guidance on this topic is fairly prescriptive, requiring sequential tests for normality and lognormality and then providing further options. Over a large site broken into many parcels, such as is the case here, the overall result can be questionable even though each individual decision is defensible. Uncritical acceptance of the results of formal parametric statistical tests can easily lead to use of different methods for estimation of EPCs in adjacent parcels. It is reasonable to ask whether contaminant distributions really alternate from normal to

lognormal to neither over short distances or whether that apparent result is simply an artifact of (non-random) sampling. It is also reasonable to ask whether a normal distribution is ever a good choice for environmental measurements given the possibility of negative (impossible) values. Use of a consistent methodology (e.g., assumption always and everywhere of lognormality) would produce a result easier to understand, more defensible (in my opinion), and less time consuming to obtain.

4. Were the values used to represent the exposure and absorption parameters used in the direct-contact exposure assessment appropriate under the evaluation criteria, specifically:

• Exposure duration for each scenario;
• Exposure frequency and area use factors for each scenario and exposure area;
• Soil ingestion rates;
• Exposure assumptions affecting dermal contact (e.g., soil adherence rates, skin surface areas assumed to contact soil or sediment); and
• Oral and dermal absorption factors.

In addressing this question, please consider the same factors listed in Question 1 (as relevant).

Generally exposure frequencies and durations appear reasonable or conservative. The selected exposure frequency for direct contact by farmers (10 days per year) is implausibly low.

EPA uses residential soil ingestion rates of 200 mg/day for children and 100 mg/day for adults. GE argues these rates are too high. Given that EPA's standard assessment practice does not consider the possibility of a geophagic child (and is therefore unlikely to be protective of such an individual) further reduction in soil ingestion rate estimates should only be considered on strong evidence. GE offers a letter by Calabrese in support of lowered values. Calabrese and his coworkers have a history of publishing multiple and inconsistent conclusions from the same datasets and of non transparent derivation of their estimates. Indeed, the existing EPA defaults are based on results published by Calabrese's group. As recently as 1995 (Environ Health Perspec 103:276-285) Stanek and Calabrese reported an estimate of 1200 mg/day for the 90^{th} percentile annual average daily soil ingestion rate for children in the U.S. Child soil ingestion rate estimates attributable to Davis et al. (Arch Environ Health, 45(2):112-122, 1989) should be viewed as at least as valid as estimates generated by Calabrese et al. Estimated 95^{th} percentiles from the Davis et al. data using silicon and aluminum as tracers exceed 200 mg/day. Reduction of the child RME value is therefore not justified.

The default dermal absorption value for PCBs (14%) used in the HHRA is based on results reported by Wester et al. (1993). The laboratory protocol employed by Wester et al. has multiple weaknesses including failure to record PCB-soil contact time prior to exposure, use of an inappropriate particle size, use of an animal model that precluded sacrifice and determination of mass balance, and use of a vertical animal posture that would permit soil

sloughing. Nevertheless it is impossible to state with certainty whether the default value is conservative or non-conservative. Since the default has been in use for some time, it should, in the interest of consistency, continue to be used until such time as a more defensible factor is available. The methodology described in a recent paper by Mayes et al. (2003) repeats some of the mistakes of Wester et al. and compounds them with additional design flaws. Most significantly, Mayes et al. applied soil in amounts that would represent many (probably 20 or more) layers. When results are expressed as percent absorbed, layering will lead to artificially low apparent absorption. Cautions regarding the effects of layering on dermal absorption can be found in EPA guidance dating back to 1992. Correction of the results reported by Mayes et al. leads to the conclusion that equivalent absorption from a monolayer would have been on the order of 80%. Hence downward adjustment of the EPA default on the basis of the Mayes et al. experiments is completely unwarranted.

Although it is taken directly from 1992 EPA guidance in which layering effects are discussed, the EPA default estimate of dermal bioavailability for dioxin (3%) cited in the ROR HHRA was not derived from experiments conducted at monolayer loading or corrected for layering. It is therefore inconsistent with EPA guidance even though it represents EPA guidance.

It should be noted that the dermal bioavailabilites cited above are 24 hour bioavailabilities. Use of a fixed dermal absorption factor expressed as a percent, although consistent with current EPA guidance, is poor practice as it ignores the time dependence of absorption (or effectively leads to the unlikely assumption that all dermal exposures last for 24 hours). Consideration of temporal dependence of dermal absorption is a refinement that has not yet been incorporated into EPA guidance related to soil contamination at Superfund sites.

5. Is the approach used to estimate a Reasonable Maximum Exposure (RME) and a Central Tendency Exposure (CTE) for the direct-contact exposure assessment appropriate under the evaluation criteria?

EPA guidance for estimation of RMEs is subject to interpretation. Some unstated number of exposure factors are to be incorporated at upper percentiles while others are held at central tendencies. This leaves ample room for production of disparate estimates of the RME by investigators using the same data. With this caveat, the general approach taken appears reasonably consistent with EPA guidance.

6. Were the uncertainties adequately characterized and expressed?

Uncertainties are discussed in qualitative terms, but the Direct Contact Exposure Assessment is a deterministic analysis. The numerical results simply do not express uncertainty. (I do not agree that estimation of two values, the RME and CTE, which are of essentially unknown statistical character, constitutes uncertainty analysis.)

7. Overall, was the approach used to estimate risk from direct contact reasonable for evaluating the baseline risk?

The overall approach is generally consistent with EPA guidance and practice. Some shortcomings of the prior screening process (see discussion above) may have resulted in premature elimination of exposure areas that should have been retained. Any such failures would have simply been carried through the direct contact assessment.

C. Phase 2 - Fish and Waterfowl Exposure Assessment

1. Were the approaches and methods used to calculate EPCs for the fish and waterfowl consumption scenarios appropriate under the evaluation criteria?

Methods used to calculate EPCs appear to be generally reasonable and consistent with EPA guidance. Missing data for some fish species in some reaches and waterfowl in Connecticut are weaknesses.

2. Were the exposure assumptions and parameters used in both the assessments of fish and waterfowl consumption appropriate under the evaluation criteria?

The discrepancy between the assumed sizes of fish (8 oz) and waterfowl (4 oz) meals requires justification.

EPA appears to have misinterpreted the consumption rates from Ebert's 1993 Maine angler study.

3. Was the basis for the selection of point estimate RME and CTE exposure parameter values appropriate under the evaluation criteria, and were they clearly described and referenced?

Exposure factor selections were adequately documented. See two comments above.

4. Were the probabilistic approaches used clearly described, and were they appropriate under the evaluation criteria?

In general I find the probabilistic approaches unsatisfactory. This is due in part to the fact that I find mixed deterministic/probabilistic approaches to be confusing (especially to the general public). EPA justifies using both deterministic and probabilistic methods by citing guidance regarding tiered approaches. While this justification may be valid, the general organization of the tiers is poor. They do not logically flow from one to the next but are intermingled. Moreover, the rationale for applying some form of probabilistic assessment only to the Fish and Waterfowl Exposure Assessment is not adequately developed. If it is the view of the authors of the ROR HHRA that fish and waterfowl pathways are the most

significant and therefore deserving of the most sophisticated treatment, then that position should be revealed through an appropriate report structure (i.e., comparable initial assessments of all pathways providing evidence of quantitative differences followed by more sophisticated treatment of the most critical pathways). It is entirely possible that such an approach would have justified differential treatment of the fish and waterfowl pathways, but the case has not been adequately made. Using probabilistic methods within just one section in the middle of the overall report and then applying multiple techniques within that one section renders the reader's job more difficult.

5. Were the distributions used in the probabilistic assessments clearly described, and were they appropriate under the evaluation criteria?

Generally adequate information was provided to understand what assumptions were used. Why point estimates were used for fish concentrations and ingestion rates in an ostensibly probabilistic analysis is not clear. Mixed analyses produce hybrid results that are very easily misinterpreted.

6. Were the uncertainties in the data and models adequately characterized and expressed?

Use of PBA to characterize uncertainty generally produces very wide bounds and provides no information regarding degree of confidence within those bounds. I believe that 2-dimensional Monte Carlo analysis is easier to implement and understand. I would much prefer to look at a set of 2-D Monte Carlo plots than Figure 8-2 in Vol. IV Appendix C.

Resolution of the question of whether Native American subsistence fishermen do or do not exist in the Housatonic watershed is necessary. A screening level assessment of exposure to a subsistence fisherman should probably be included in any case.

7. Were variability and uncertainty in the risk estimates adequately characterized and expressed?

Use of point estimates within the Monte Carlo assessment that was supposed to characterize variability would be expected to produce an underestimate of population variability.

Graphical presentation in Volume IV, Section 6 would have been improved in many cases by use of log scales on the x-axis.

8. Overall, was the approach used to assess risk from consumption of fish and waterfowl and other wild food items reasonable for evaluating the baseline risk?

Because multiple approaches were used, it is difficult to summarize. Some inputs to the deterministic assessment should be reconsidered and point estimates should not have been

applied in the probabilistic assessment. I am unable to quantitatively assess the effect of needed changes. However, I note that the alternative analysis of fish consumption in Reaches 5 and 6 provided in comments to the panel by AMEC on behalf of GE produces a 95^{th} percentile risk approaching 10^{-3} even after downward adjustment of multiple exposure factors. This might be viewed as evidence of consensus regarding significant risk via fish consumption in at least a portion of the Rest of River study area.

D. Phase 2 – Agricultural Exposures

1. Were the exposure scenarios evaluated appropriate and reasonable for current and reasonably foreseeable future use of the floodplain?

Assessment of exposures at the fixed and somewhat arbitrarily chosen soil contamination levels of 0.5 and 2.0 mg/kg is inconsistent with assessment based on actual contamination levels in the rest of the document. The 2.0 mg/kg value is based on an agreed upon cleanup level, but it represents a hypothetical future state and therefore leads to summary/graphical results that are not directly comparable to results generated for the other pathways.

Adequacy of the treatment of future land use is unclear. Tabular description of past and present land use showing trends (by parcel or reach, not land use) would improve the presentation.

2. Were the approaches used to estimate transfer of COPCs from soil to plants appropriate under the evaluation criteria?

Use of site specific data is generally good practice, but in this case so little data are available that the overall effort must be considered very uncertain. In the absence of formal (quantitative) uncertainty analysis, this weakness is effectively hidden.

Use of data from washed beets and turnips does not appear reasonable.

3. Were the approaches used to estimate the bioaccumulation of COPCs in animal tissue appropriate under the evaluation criteria?

Did not evaluate.

4. Were the exposure assumptions and parameter values appropriate under the evaluation criteria?

Use of a 75^{th} percentile consumption rate for RME estimation is inconsistent with use of 90-95^{th} percentiles elsewhere in the document.

5. Was the basis for selection of values clearly described and referenced?

Generally assumptions were explained. Justification for those assumptions was in some cases less obvious.

6. Is the approach used to estimate the RME and CTE appropriate under the evaluation criteria?

Since exposures were calculated using the chosen EPC values of 0.5 or 2.0 mg/kg rather than measured/estimated values, RMEs and CTEs calculated here are not directly comparable to RMEs and CTEs calculated elsewhere in the document.

7. Were the uncertainties in assessment adequately characterized and expressed?

As in the Direct Contact exposure case, uncertainties are discussed in qualitative terms, but the Agricultural Exposure Assessment is a deterministic analysis. The numerical results simply do not express uncertainty.

8. Overall, was the approach used to assess risk from consumption of agricultural products and other wild food items reasonable for evaluating the baseline risk?

Given the shortage of data for some key factors such as transfer coefficients, this section would benefit greatly from formal uncertainty analysis.

E. Phase 2 – Integrated Risk Evaluation

1. Were the bases for the toxicity assessment adequately described including the cancer slope factors, reference doses, and calculations of TEQ?

The toxicity assessment for total PCBs generally followed longstanding EPA practice and could be readily understood. Treatment of TEQ risks was much less clear. I did not make an attempt evaluate the latter and defer to reviewers with more expertise in toxicology.

2. Did the risk characterization describe the methods and risk summary at an adequate and appropriate level of detail?

Bar charts that show ranges from lowest CTE to highest RME are not a good communication tool. CTE and RME ranges should be shown separately.

3. Were the potential risks associated with exposure to a combination of pathways and COPCs (direct contact, fish and waterfowl consumption, and agricultural product consumption) adequately characterized?

Aggregate exposures were generally not considered. Aggregation was essentially dismissed by comparing fish eating or waterfowl eating to soil eating risks for fishermen and hunters. No attempt was made to estimate, for instance, how many different Direct Contact pathways a non hunting/fishing individual might experience. The panel was not briefed on the manner in which cleanup decisions will ultimately be made. Conceivably, evaluation of individual direct contact pathways in isolation could lead to a no action decision that is inconsistent with EPA policy.

4. Were the uncertainties associated with both cancer and non-cancer health effects adequately characterized and expressed?

Generally no. By policy, EPA does not consider probabilistic treatment of uncertainty in toxicological dose response factors, but does prescribe point values to be used. While derivation of those values includes adjustment (in a conservative direction) for uncertainty, the resulting point estimate does not express either variability in population response or uncertainty due to ignorance. Discussion buried in text is not an adequate substitute for numerical and graphical representation of confidence bounds.

F. General

1. Were the EPA toxicity approaches and values (e.g. IRIS and HEAST) used for the COPCs applied appropriately under the evaluation criteria?

Consideration of toxicity in the ROR HHRA appears to be consistent with current EPA guidance. Note limitations cited above.

2. Were the important assumptions for estimation of dose (i.e., toxicity and exposure) and risk identified?

Yes, assumptions were generally stated.

3. Were the calculations of carcinogenic and non-carcinogenic risks performed properly and consistent with EPA guidance?

Risk calculation for total PCBs in the ROR HHRA appears to be consistent with current EPA guidance. I did not review the estimation of carcinogenic risks attributable to TEQs and take no position on that aspect of the work.

4. Were the significant uncertainties inherent in the risk evaluation properly addressed and characterized? If not, please identify those that were not properly addressed or characterized and how they should be addressed in the HHRA.

EPA's standard practice in conducting Superfund risk assessments remains primarily deterministic in character. With the exception of the Fish and Waterfowl Exposure Assessment that is the case here. By definition a deterministic approach cannot adequately address uncertainty and variability which are statistical concepts. Paired CTEs and RMEs, which represent point estimates at unknown and undoubtedly variable percentiles of the overall population distribution, provide little or no characterization of uncertainty. If one accepts that EPA's deterministic risk assessment methods lead not to predictions of absolute risk, but rather to estimates of relative risk that permit consistent decisions to be made, those methods can still be useful. However, that fact is often not adequately expressed in risk assessment documents. In fairness, implementation of fully probabilistic analyses of risks at Superfund sites remains problematic due to inadequate input data and failure to conduct such an analysis is not in conflict with current EPA guidance.

The ROR HHRA should include at least a screening level assessment of exposures to infants via breast feeding.

5. To the best of the Panel's knowledge, have relevant peer-reviewed studies that support, are directly relevant to, or fail to support any estimate of risk been identified and considered, and has an appropriate methodology been used to reconcile inconsistencies in the scientific data?

Note discussion of child soil ingestion estimates attributable to Davis et al (1989) above.

6. To the best of the Panel's knowledge, is there other pertinent information available that was not considered in the HHRA? If so, please 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 HHRA.

Regional biomonitoring results are suggestive of a non-occupational contribution to elevated serum PCB levels in local populations. The panel was provided a 1997 MA DPH report (Housatonic River Area PCB Exposure Assessment Study – Final Report) describing blood serum levels observed in two study groups designated "Exposure Prevalence" (n=69) and "Volunteer" (n=79). The panel was not provided raw data from this report, so only limited interpretation is possible. However in both study groups, persons scoring in the top quartile on a questionnaire regarding behaviors that might reasonably lead to environmental PCB exposure (i.e., persons most likely to receive non-occupational exposure) had numerically higher median blood PCB levels than did persons who scored in the lower three quartiles. (See Figures 13 and 14 in the MA DPH report.) This effect was observed in persons with and without occupational exposures who were appropriately segregated. (So the effect occurs in 4 of 4 cases.) Behaviors queried in the survey were directly relevant to pathways

considered in the ROR HHRA (e.g., fishing, flood plain activity, etc.). It would be reasonable for EPA to examine these apparent differences for statistical significance and compare implied doses to exposure predictions generated in the ROR HHRA. Observed blood PCB levels exhibit strong dependence on age. If those scoring in the upper quartiles on the behavioral survey were older than those scoring in the lower quartiles, the apparent environmental contribution to body burden may be artifactual. Differences could also be attributable to exposure pathways other than those considered in the ROR HHRA. For instance, elevated body burdens might be expected in family members of occupationally exposed persons or in persons living in housing contaminated by prior occupants who were occupationally exposed. Those persons might have coincidentally scored higher on the behavior survey. (This would not explain higher blood levels at higher scores within the occupationally exposed groups however.) Nevertheless, it is useful to extract dose estimates from the ROR HHRA and compare them to estimated inputs that would be required to maintain observed body burdens.

Differences in median blood levels between highest quartile behavioral score individuals and lower score groups appear to be 2-4 ppb. Making rough assumptions about body weight, body fat, blood-fat partitioning, and physiological half-lives of PCBs, it is possible to make back-of-the-envelope estimates of steady state inputs required to account for incremental increases in that range. Resulting doses are roughly those that are projected in the ROR HHRA to present lifetime excess carcinogenic risk on the order of 10^{-5}. In Figure 7-1 (Vol. I) all pathways that have upper bounds in the 10^{-4}-10^{-5} range are rendered plausible in that some individuals in the population appear to have body burdens elevated (over within-study occupational or non-occupational background) by amounts that would be produced by the projected exposures if they were actually occurring now. The highest risk projections shown in Figure 7-1 (Vol. I) do imply PCB doses that would lead to very high blood levels (~ 1 ppm). Failure to identify any such person in the MA DPH biomonitoring studies could indicate that those projections are very conservative or that the sample size was too small to find them. But most projected PCB doses underlying the risk estimates presented in the ROR HHRA cannot be dismissed as unduly conservative on grounds of lack of consistency with observed body burdens.

7. With respect to the conclusions in the HHRA report:

• Are the conclusions (risk characterization) supported by the information presented in the other sections of the report?

Generally the qualitative arguments leading to the risk estimates can be followed through the documents. However, the sheer complexity of the case precludes presentation of all intermediate numerical steps and therefore renders actual reproduction of all numerical estimates impossible.

• Do the conclusions (risk characterization) objectively and reasonably characterize potential current and reasonably foreseeable future risks to human health in the Rest of River area?

Reasonable characterization of risk must be evaluated in context. The ROR HHRA is primarily a deterministic risk assessment. While this is consistent with past EPA practice and current EPA guidance, I view deterministic assessments not as true characterizations (predictions) of expected outcomes, but rather as socially negotiated conventions that enable decisions to be made. Ultimately they lead to action at some sites and no action at others. Generally it is reasonable to expect that sites will be appropriately ranked on a relative scale. Any assumption that absolute risk is well defined by a deterministic risk assessment is questionable at best. Nevertheless it appears that there is substantial reason to expect, under current conditions, elevated human exposure to PCBs as a direct consequence of contamination of the Housatonic River and its flood plain.

This question also specifically invites consideration of the foreseeable future. However the ROR HHRA does not address the environmental fate of PCBs (or other COPCs). Some increased downstream transport of PCBs might reasonably be anticipated, especially if a dam holding significant mass of PCB-contaminated sediment were to fail or deliberately be removed or if bridge construction/repair were to disturb sediments. A large flood event might also transport relatively high concentration sediments to areas currently delineated as flood plain or extend the 1 ppm isopleth. These kinds of events might lead to higher PCB exposures among upper reach or downstream populations than would be currently estimated. Conversely degradation or volatilization loss of PCBs (if significant) could lead to reduced exposures. The ROR HHRA assumes static environmental conditions and in this sense does not adequately characterize possible future conditions. Consideration of future land use is also marginal. Generally the ROR HHRA assumes that land uses will not change or will change in a way that reduces potential for exposure (e.g., less farming). The possibility that existing land use represents an artificially depressed state as a consequence of public concern over contamination and that a general decline in concern with time (issue fatigue) will lead ultimately to less cautious uses does not appear to have been considered. Discussion of changes in land uses that might lead to increased exposures would inform more appropriate discussion of need, if any, for statutory restriction on land use.