Dr. Roland B. Hemmett
U.S.EPA, Region 2
Edison, New Jersey 08837

Dear Dr. Hemmett,

I have reviewed the questions which you sent to Wayne Munns, and the appropriate supporting materials. As Dr. Munns was not able to respond before to your request before he left the country, and will not return until after the deadline, I will respond to those of the questions for which I have some expertise. Additionally, I have consulted Drs. Richard Pruell and James Lake, both of this laboratory, for input on several questions. Our answers can be found below. The responses are primarily from me, unless I have indicated otherwise. We have not responded to questions 2,7,11,14, or 15.

Sincerely,
Walter J. Berry, Ph.D.
Research Biologist

cc. W. Munns
N. Rubinstein
J. Lake
R. Pruell

Attachment

Response to final peer review questions.

Walter Berry, James Lake, Richard Pruell, U.S.EPA/AED

1. The framework does seem to be scientifically appropriate. For the most part it does seems to represent the "state of the art." Comments on some of the individual components of the framework may be found below.

3. All eight factors appear relevant to a degree. The application of the seventh factor (the extent of toxicity" seems a bit problematical, because if the sediments are acutely toxic there may be no need to do bioaccumulation testing (because the sediment will fail based on toxicity alone). At the same time, I imagine most resource managers would be more comfortable with allowing a sediment causing a small amount of bioaccumulation to pass, if the sediment was not toxic. Similarly, most managers would probably be more comfortable failing a sediment, based on bioaccumulation, if there was also some toxicity associated with the sediment.

It is not clear how these factors could be put into a quantitative system. Ostensibly such a system, which might involve scoring and weighting of the individual factors, would be less arbitrary. However it would probably serve only to put the arbitrariness up front, instead of when the assessment is made.

It seems important to do an integrated effects evaluation of the bioaccumulation result because most of the other evaluations relate to single compounds or single classes of compounds, and do not relate results with those in the toxicity tests.

4. A. The matrix values seem to be a suitable tool for use as a part of an evaluation, although they would probably not be suitable if other methods were not employed (But see 4B). Exceedances of these values may serve as an appropriate screen, but obviously they can not be used to calculate risk. If the matrix values are used, It might be argued that a grand mean is too low. However, this must be balanced by the fact that the exposure in the test is only 28 days, and that an arithmetic mean might be particulary susceptible to being skewed by "hot spot" values. Depending on the statistical distribution of the data, a geometric mean, or the median, might be more appropriate.

4. B. It is not clear to me why the matrix values would continue to be used, now that the background tissue values are available. The background values use newer analytical techniques and come from a more defined database.

5. A. Dr. Pruell felt that there are not enough data to support or refute any of the values proposed for dioxin at this time.

5. B. Dr. Pruell also felt there was a need to use a risk approach to the assessment of dioxin in tissues, and that considerations as to the amount of material to be dumped, and the characteristics of the dump site may be more important in this case than the tissue value chosen for dioxin.

6. A. To the extent that FDA action levels are derived based on human health considerations they would seem to be useful. To the extent that they were driven by considerations of analytical capability and background concentration, or economic considerations, they probably are not useful because analytical techniques have improved, and background concentrations have decreased on at least some of these compounds.

6. B. We do not think that FDA limits are so low that they might cause a sediment which probably poses no risk to fail, so if used in a screening mode, we do not see why omitting them would improve the assessment. It might be argued that they give a false sense of security because they are too high. On the other hand, some might be concerned if the "FDA" limits are no longer used, because they are associated with food. This is probably more a political issue than a scientific one. If there are more recent values in use or being considered by EPA for some of these compounds, the FDA values should not be used to exclude these values.

8. The CBR-type approach seems to be the most reasonable one to use with narcotic chemicals. Even though the WQCTL approach seems valid, it was not the approach chosen by the Office of Water (OW) for development of its Equilibrium-Partitioning Derived Sediment Guideline (ESG) for PAHs. The approach OW chose is summarized in a briefing document for a presentation to the EPA Science Advisory Board entitled, "Assessing the Toxicity and Bioavailability of PAH Mixtures in Sediments". May 13, 1997. I would encourage you to examine this approach, which explicitly looks at all of the PAHs together, using an additivity model. I will not comment on the WQCTL approach except to say that the CBR model uses molar units of PAH because those are the appropriate units for comparison. Although it may be appropriate at some point in the analysis to convert to weight units for PAH, it is not appropriate to say that a 400 ppb dose of napthalene would elicit the same effect as a 400 ppb dose of fluorene (as stated in the MOR), nor is it strictly appropriate to add the wet or dry weight concentrations together (as stated in the MOR).

9. Dr. Lake did not think that total PCBs should be based upon a limited subset of congeners. Organisms can substantially alter PCB distributions relative to those present in Aroclors or in sediments, and by measuring only a limited subset substantial errors can be introduced. He thought that the methods for quantitation of all, or almost all, congeners exist, and these procedures are not much more difficult than those already in use for determining the 22 congeners currently measured. Dr. Pruell felt that, although some information might be lost by looking at only a subset of the PCBs, the error would be less than a factor of 2-3, an error that seems easily within the level of variability in many of the other assumptions required to assess the risk of PCBs. He felt that the small gain from doing all of the congeners was probably not worth the added expense.

10. Dr. Pruell felt that the multipliers were probably the best available, and knew of no others.

12. Dr. Lake felt that PAHs pose a different problem than PCBs with regard to quantitation. Presently, there isn't a good way to effectively quantitate all PAHs (including alkyl homologues) present in extracts. The only alternative is to select a subset. Dr. Pruell agreed, and added that eventually we will need to get a better understanding of the effects of the other PAHs (as well as the saturated organic

compounds).

13. Dr. Pruell felt that a trophic transfer coefficient of one may not be appropriate for the metals. It is probably too low for mercury, and too high for the other metals. Further, Dr. Pruell felt that there were good transfer coefficients available for many of the organics. Much of the relevant work has been done at EPA's laboratory in Duluth, Minnesota, by Phil Cook and others.

16. The Squibb et al. (1991) reports seems appropriate for identifying contaminants of concern. We would not delete or add any compounds from the list currently tested, but would recommend the use of approaches which allow chemicals to be summed within a chemical class (e.g. narcosis).

17. Synergistic effects are poorly understood, except within classes of compounds (e.g. narcosis within narcotic compounds). We do not see how they could be included. Dr. Pruell thought that antagonism was probably more commonly reported than synergy, but that in the absence of additional data additivity was probably the most reasonable approach.

18. The basis for the "10X" criterion is not stated, so it is difficult to assess its suitability. If the "10X" criterion is derived from an understanding of the sediment-to-sediment variability in bioaccumulation tests, then it might be appropriate. However, Dr. Pruell could not see where the 10X number came from, and felt that it might be appropriate in some samples near the detection limit, but that it may be too high in samples where a compound is present at concentrations well above the detection limit.

19. An even weighting seems appropriate.

Without knowing the distribution of the data is difficult to know which measure of central tendency is best, as was true with the matrix values. An arithmetic mean might be particulary susceptible to being skewed by "hot spot" values. Depending on the statistical distribution of the data, a geometric mean, or the median, might be more appropriate.

The assumptions on page 14 seem reasonable.

20. It seems appropriate to use background concentrations from organisms in the area around the HARS, but not from organisms directly in the HARS, if the assumption is that the HARS is degraded. If the goal was only to prevent further degradation it might be appropriate to use organisms from within the HARS as well.