SCIENTIFIC PEER REVIEW OF THE EPA REGION 2/CENAN FRAMEWORK FOR EVALUATING DREDGED MATERIAL FOR PROPOSED PLACEMENT AT THE HARS.

BY
PROFESSOR KEITH R. SOLOMON
CENTRE FOR TOXICOLOGY AND DEPARTMENT OF ENVIRONMENTAL BIOLOGY
UNIVERSITY OF GUELPH, GUELPH, ON N1G 2W1

WITH

DR P. SIBLEY
CENTRE FOR TOXICOLOGY
UNIVERSITY OF GUELPH, GUELPH, ON N1G 2W1

August 24, 1998

TABLE OF CONTENTS

1 INTRODUCTION 1

2 GOALS OF THE PEER REVIEW 1
2.1 BACKGROUND 1
2.2 TASKS 2

3 GENERAL COMMENTS 2

4 RESPONSES TO SPECIFIC QUESTIONS 3
FRAMEWORK 3
BENCHMARK AND RISK EVALUATION VALUES 6
CALCULATIONS 14
GENERAL 16

5 REFERENCES 18

1 INTRODUCTION

This peer review was undertaken in accordance with the goals outlined in Section 2 below: The peer review consisted of a scientific assessment of the **MEMO FOR THE RECORD** on the **SUBJECT**: Review of Compliance with the Testing Requirements of 40 CFR 227.6 and 227.27, and the Site Designation Provisions of 40 CFR 228.15 for the Project XXXX, New York, New York. The review makes general comments, and then responds to specific questions put to the reviewers. For ease of reading, the specific questions in Section 5 are highlighted with a horizontal blue lines and have a pale blue background. Responses are interspersed between the questions.

2 GOALS OF THE PEER REVIEW

The goals of the peer review are taken directly from the guidance to the peer reviewers and are appended below.

2.1 BACKGROUND

The August 29, 1997 Final Rule, Simultaneous De-designation and Termination of the Mud Dump Site and Designation of the Historic Area Remediation Site, specifies that the historic area remediation site (HARS) will be remediated with uncontaminated dredged material (i.e., dredged material that meets current Category I standards and will not cause significant undesirable effects including though bioaccumulation; hereinafter referred to as *Remediation Material*). The rule further specifies that the HARS will be managed so as to reduce impacts within the Priority Remediation Area (PRA) to acceptable levels in accordance with 40 CFR 228.11. Placement of dredged material within the PRA is restricted to Remediation Material. This material will not cause significant undesirable effects, including through bioaccumulation or unacceptable toxicity in accordance with 40 CFR 227.6.

Evaluation of proposed dredged material regarding unacceptable toxicity is clearly defined in the Green Book as statistical criteria which require no interpretation. Evaluation regarding significant undesirable effects including through bioaccumulation requires assessment of chemical analyses of tissue from

28-day bioaccumulation tests. There are no specific regulatory criteria for this evaluation; however there are existing regional guideline values that have been developed and used, by the U.S. Environmental Protection Agency (EPA) Region 2 and the U.S. Army Corps of Engineers New York District, to evaluate the constituents in accordance with 227.6.

This peer review charge is to assess whether the testing evaluation process is adequate to properly determine whether a tested sediment is suitable for Remediation Material as defined. Your review should focus on the framework for evaluation of bioaccumulation data and guideline values used; it should not deal with on toxicity/mortality testing. Please bear in mind that the testing evaluation applies to risks pertaining to ocean placement of the sediment, and not to risks pertaining to other alternatives such as leaving the sediment in place.

2.2 TASKS

This charge is in the form of questions on critical aspects of the evaluation framework. General references are cited in each charge question to aid in finding the issue in question. Note that these are general guiding referrals and should not be considered the only review item for those specific issues. Please answers the assigned questions as directly as possible, given the provided materials and your own expertise. If you are unable to answer a particular question on the basis of the provided materials, please inform us of information needed to answer the question. Also, keep in mind that there are additional environmental data resources and test data pertaining to the New York Bight available in EPA Region 2, if they are needed.

3 GENERAL COMMENTS

In general, the EPA Region 2/CENAN framework for evaluating dredged material for proposed placement at the HARS followed a framework consistent with those commonly used in environmental and human health risk assessments. The approach was tiered to focus issues on key points and the

analysis of the data was correctly done. Given the background information and the data presented in the Memo, the conclusions are completely justified.

4 RESPONSES TO SPECIFIC QUESTIONS

FRAMEWORK

1	Is the EPA Region2/CENAN Framework for evaluating bioaccumulation results scientifically appropriate for determining the suitability of dredged material as Remediation Material? If not, describe deficiencies. (Please see Region2/CENAN joint evaluation memorandum, Figure 1)

RESPONSE

The framework for evaluating the results of the bioaccumulation results (as described on pages 6-8 in the memorandum) is a reasonable approach to hazard assessment and is valid. The only potential problem that could result from the decision tree approach is when the concentrations in the reference sediment are high enough that criteria, such as the FDA levels, were exceeded. This would be the case if an inappropriate reference sediment were used. Inspection of the data in Table 1 revealed that this was not the case. In addition, the reference sediment was collected from an appropriate location.

2	Which of the risk-based values derived constitute *true* conservative estimates of risk levels (i.e., exceeding the value should be interpreted as sufficient cause to conclude that significant undesirable effects may result through bioaccumulation)? Which of the risk-based values derived constitute conservative screening values (i.e., test tissue concentrations below the value can confidently be interpreted to pose no risk of significant undesirable effects and exceeding should be further evaluated before the probability of significant undesirable effects can be assessed)? How can the *true* risk levels be calculated for those compounds which you believe only to have screening values? How should test concentrations be compared to risk-based levels to determine whether they are exceeded.

RESPONSE

The term risk is often used in the incorrect context. Use of the term “risk” implies that the likelihood of something happening is known or has been estimated. Properly, risk should always be expressed as a probability. Comparison of a concentration (in biota, or in a matrix) to a reference concentration or a

criterion concentrations is an assessment of hazard. Hazards can either be present or not be present, as the one concentration is either greater than or less than the other. Traditionally, hazard quotients (one concentration divided by another) have been used in the early tiers of risk assessment to determine whether further and more detailed risk assessment is needed. The criteria or standards used to calculate reference values for hazard quotients are usually based on relatively conservative numbers. For example, the procedures to calculate water quality criteria use a number of conservative assumptions (Stephan et al. 1985). The Final Acute Value criterion is based on the more sensitive organisms (5^{th} centile of the genus mean acute values) and additional conservatism is added in the calculation of the Final Chronic Value. The reason for these conservative approaches is that the criteria are designed to be protective of almost all organisms, most of the time. The criteria are designed to apply in a variety of situations, some where for physical or biological reasons, more sensitive organisms may be present, while for other they may not. The criteria are thus protective, not predictive. The use of these hazard quotients to assess "risk" is therefore conservative.

Used in the proper way, the hazard quotient can be used to decide whether a hazard exists or not. If it does not exist, the situation is unlikely to present a significant hazard and no further risk assessment is needed. However, the obverse, that is, the exceedence of the criteria, does not mean that a significant risk exists, it merely means that further work is necessary to better quantify the risks. Given that none of the criteria were exceeded in this particular risk assessment suggests that further detailed risk assessments are not necessary as the criteria on which the hazard quotient was based are conservative. Had some of these values exceeded the established criteria, other approaches to risk assessment such as those using probabilistic techniques (Klaine et al. 1996; Parkhurst et al. 1995; Solomon et al. 1996) could have been used, provided that sufficient data were available to adequately describe the range of susceptibility of organisms and the spatial and temporal variation of the exposure or body concentrations.

This reviewer is not suggesting that a probabilistic risk assessment be carried out in this particular case but rather that this may be another way of conducting these assessments once the probabilistic techniques have been refined and the appropriate data collected.

3 In conducting the integrated effects evaluation using the types of data provided by the applicant, which of the eight factors for LPC compliance listed in the Green Book are appropriate and relevant? How can a quantitative/strategic framework be established to evaluate tissue data for those factors? Considering that comparison to regional Matrix values and site-specific risk values represent case-specific evaluations, is it necessary to conduct the integrated effects evaluation of the bioaccumulation results? (Please see Reference No. 61, page 6-6)

RESPONSE

The eight compliance factors in the "Green Book" (USEPA 1991) are all reasonable but some are more biologically relevant than others. A discussion of this is summarized in the table below:

Green Book Criterion	Biological relevance	Useful-ness*
Number of species from dredged material in which bioaccumulation exceeds reference (statistical test).	Based on difference from a reference material. No toxicological relevance assessed. Incorrect choice of reference material could confound the results.	T
Number of bioaccumulated contaminants from dredged site in which exceed reference site values (statistical test).	As above, based on difference from a reference material. No toxicological relevance assessed. Incorrect choice of reference material could confound the results.	T
Magnitude by which bioaccumulation from dredged material exceeds that from reference.	More useful as it is a continuous variable, however, the response of organisms to increasing concentration (concentration response) would need to be factored in as well.	TT
Toxicological importance of contaminants from dredged site exceeding those from reference site.	Again, this is based on difference from a reference material. No toxicological relevance assessed and the importance of these contaminants is judgemental. Incorrect choice of reference material could confound the results.	T
Phylogenetic diversity of contaminated species exceeds that from reference site.	Phylogenetic diversity may not be relevant to ecological importance or function in the ecosystem, however, this may be an indicator of greater potential for entry to food chain. Diversity may be affected by physical factors such as particle size.	TT

Green Book Criterion	Biological relevance	Useful-ness*
Propensity for contaminants with statistically significant bioaccumulation to biomagnify in aquatic food chain.	Biomagnification usually only occurs with persistent and lipid soluble substances. These may have a greater impact in organisms higher on the food chain (as has been demonstrated historically) and this is judged to more useful.	**T T**
Magnitude of toxicity and phylogenetic diversity of organisms showing greater mortality in dredged material.	A good effect-based criterion that is related to response of organisms. It may, however, be confounded if incorrect matching of test and reference sediment is used. Some organisms will not thrive and "die" if sediment physical characteristics are not appropriate.	**T T T T**
Magnitude by which contaminants whose bioaccumulation from dredged site exceeds that in organisms near the proposed site.	Some usefulness but subject to confounding from poor choice of nearby sites. A good margin of safety may exist at both sites despite the differences.	**T**

* the more useful, the more **T**s

The response of the organisms at the site will, to a degree, integrate the effects evaluation. Other types of toxicological integrators (TEFs and TEQs) are less well developed. If site-specific values are used, an attempt to should be made to integrate the effects evaluation of the bioaccumulation results, however, the biological responses highlighted above should be given higher credence in the assessment. In this regard, physical properties of sediments may be more important than chemical properties. Some sediments are unsuitable substrates for colonization by some oganisms and, absence of these organisms does not mean an adverse toxic effect. Choice of the wrong sediment as a reference could result in false positives (for toxicity). In the assessment being reviewed here, this was not the situation.

BENCHMARK AND RISK EVALUATION VALUES

4 Regional Matrix Values

A Are the Matrix values suitable for determining the suitability for placement at the HARS as Remediation Material?

RESPONSE

The matrix values are generally judged suitable for determining the suitability for placement as remediation material. Values for dioxin TEFs (FDA and USEPA) are based on mammalian studies.

They are thus most suitable for assessing risk to humans (and other mammals). For assessing risks to fish, TEFs based on data from fish may be more useful (Parrott et al. 1995). However, given the observed concentrations, this difference was not judged to be significant.

	B	Regional Matrix values were developed in 1981 by compiling available field data for mercury, cadmium, PCBs, and total DDTs. Were these values derived appropriately for their intended use? Based on current data sets and scientific literature, are these 1981 values suitable for predicting the significant undesirable effect due to bioaccumulation? (Please see Reference No. 57) If not, identify more current references, data sets, and/or actual chemical specific values that would be more appropriate.

RESPONSE

The relevant decision guideline limits for mercury, cadmium, PCBs and total DDT were reviewed and were, in all cases, judged to be appropriate. For cadmium, total PCBs, and mercury, these decision guidelines were all below guideline levels developed in other jurisdictions and which incorporated appropriate safety factors. Based on the arguments presented for the decision guideline value for DDT, this value is also judged appropriate. Although this reviewer is aware of some more modern studies on DDT (such as enhances breakdown in marine sediments), the results of these would not justify more conservative decision criteria values.

5 Regional Dioxin Values

	A	Currently, the presence of 2,3,7,8-TCDD at a detectable concentration (i.e., greater than or equal to one part per trillion (pptr)) in tissues of organisms exposed to dredged material precludes its classification as Category I (hence Remediation Material); presence of the remaining dioxin/furan congeners, at concentrations of TEQs equal to or greater than 4.5 pptr, results in a similar conclusion. When 28-day tissue concentrations exceed these values, is there sufficient cause to conclude that placement of the material is not suitable as HARS Remediation Material? If not, what levels indicate sufficient cause for this conclusion? (Please see Reference No. 89)

RESPONSE

The 1 ppt criterion for TCDD and the 4.5 ppt criterion value for TEQ of the dioxins and furans other than TCDD is based on the use of a number of safety factors and conservative assumptions. It is well known that criteria for dioxin vary widely from one jurisdiction to another and even between agencies in the same country. The EPA criterion is one of the most conservative while that of the FDA (20 and 50 ppt) is in the middle of the range. Given that trophic transfers are not unity, values similar to those suggested by the FDA would be more appropriate.

B	Are dioxin values suitable for predicting the significant undesirable effects due to bioaccumulation? If not, should these values be based on a risk analysis paradigm in which the size of the human population subgroup potentially exposed through intentional behavior is compared to the size of the general population in the EPA? Since the primary route of exposure is through consumption of fish and shellfish, should the variability in potential exposure due to differences in fishing behavior (e.g., target species, seasonal preferences) be incorporated in the risk paradigm? How would a benchmark protective of human health compare to benchmarks determined using an ecological risk analysis paradigm for resident fish and piscivorous wildlife?

RESPONSE

Dioxin values are suitable for assessing the hazards resulting from bioaccumulation (with the above qualifiers taken into consideration). However, for risk assessment purposes, the likelihood of exposure in the potentially exposed population should be considered. The likelihood of consumption of contaminated seafood should incorporate seasonal and target species variability as well as the likelihood that fish will be obtained from other regions that may be less contaminated (if this is the case). If these factors are considered, exposures will normally be reduced, thus further adding conservatism to the risk assessment. Human health risk assessment is normally aimed at protection of the individual, and, because of this, usually incorporates many conservatisms. Ecological risk assessment is focused on endpoints at the population level rather than the individual. Thus, risks to fish and picivorous wildlife would be assessed differently from those to humans. Criteria based on human consumption would be expected to be protective of wildlife.

6 FDA Action Levels (Please see Reference No.61, Sec. 6.3)

A Are FDA Action Levels useful as upper limit human health benchmarks?

RESPONSE

As discussed above, FDA action levels for the protection of human health are based on protection of the individual and embody a number of conservative assumptions. They are judged entirely appropriate for the protection of human health. Although the FDA does not consider environmental effects (and some substances may be more toxic to invertebrates and fish than to mammals) the conservative assumptions used in the setting of FDA action levels will likely be protective of fish and shellfish and the function of their populations in the environment.

B Would the evaluation be improved by omitting comparison of tissue results to FDA Action Levels?

RESPONSE

In the opinion of this reviewer, the assessment would not be improved by omission of the tissue concentrations to FDA action levels.

7 Human Health Risk, Cancer and Noncancer

A Are the risk values suitable for determining the suitability for placement at the HARS as Remediation Material? If there are better alternatives for human risk, specifically what are they?

RESPONSE

As discussed above, the FDA human food consumption guideline values are conservative assumptions used for the protection of individual humans. They do not consider the likelihood of consumption of contaminated seafood and do not usually incorporate seasonal, catch site, and target species variability. If these factors are considered, exposures would normally be reduced, thus further adding conservatism to the risk assessment. Probabilistic approaches to assess the likelihood of consumption would be more appropriate.

	B	Benthic tissue levels for cancer protection were derived using assumptions focused on attaining a cancer protection at the 10^{-4} risk level. Is this risk appropriate for a determination of ocean placement of Remediation Material? (Please see Region2/CENAN joint evaluation memorandum, Appendix for Table 1, Page A-4, A-5)

RESPONSE

The use of the multistage linear model for extrapolation of risks from laboratory animal studies to humans is very conservative. For one, it does not consider the presence of threshold of toxicity (carcinogenicity). Biologically, all effects likely have thresholds, it is just that these thresholds cannot easily be demonstrated experimentally. Repair mechanisms for many of the cancer-causing mutational events exist and function to repair damage from natural mutational events. These natural mutational events are usually far more numerous than those caused by low exposures to synthetic chemicals. Not all species of fish or shellfish would necessarily be consumed by humans, thus adding further conservatism to the assessment. The use of a 10^{-4} cancer risk estimate is therefore judged to be appropriately conservative for the purposes of ocean placement.

	C	Benthic tissue levels for noncancer protection were derived using Reference Dose (RfD) of several organic and inorganic contaminants for the protection of human health. Are these values appropriately and consistently derived? Is the whole body/fillet conversion factor of 1.35 an appropriate factor for all of the contaminants considered if human exposure is assumed to be primarily via consumption of the fillet portion of the fish? (Please see Region2/CENAN joint evaluation memorandum, Appendix for Table 1, Attachments B and C) If not, what factors would be appropriate? For the lead noncancer value, since there is no RFD for lead the value was derived differently than the other metals. Was the value derived appropriately? (Please see Reference No. 88)

RESPONSE

The methods used to determine benthic tissue levels for the protection of human health were judged to be appropriate and consistently derived. The whole-body fillet conversion factor of 1.35 is judged to be slightly conservative (based on this reviewers experience with organochlorine concentrations in fish tissues). Fat is consumed to produce energy in fish muscle (fillet) and lipid concentrations (and associated lipid-soluble materials) are usually significantly lower than in other tissues (in our studies

muscle had less than 1% fat while the carcase had between 3.7 and 5.6% fat). These conversion factors and the Gobas trophic transfer model are judged appropriate for determining possible fish exposure concentrations.

The RFD for lead was derived from exposure concentrations appropriate for the protection of children, the most sensitive human life stage for this element. The RFD considered exposure via other routes and is judged to be appropriate.

D	Are the risk values suitable for predicting the significant undesirable effects due to bioaccumulation? Since the primary route of exposure is through consumption of fish and shellfish, should the variability in potential exposure due to differences in fishing behavior (e.g., target species, seasonal preferences) be incorporated in the risk paradigm?

RESPONSE

As discussed above, human food consumption guideline values are conservative assumptions used for the protection of individual humans. They do not consider the likelihood of consumption of contaminated seafood and do not usually incorporate seasonal, catch site, and target species variability. If these factors are considered, exposures would normally be reduced, thus further adding conservatism to the risk assessment. Probabilistic approaches to assess the likelihood of consumption would be more appropriate.

8 Ecological Risk

A	Ecological effects benchmarks include the Water Quality Criteria Tissue Level (WQCTL), Critical Body Residue (CBR) associated with narcotic responses, and certain mutagenic/teratogenic effects. Is it valid to use the CBR effect end point for evaluating significant undesirable effect? Are there other ecological end points that should be used to measure ecological risk that are protective of marine benthic and fish life via trophic transfer, particularly for PAHs? If so, identify. With regard to a narcotic effect for chlorinated organic compounds, should an additive approach be considered to include the contribution of chlorinated hydrocarbons against this narcotic (CBR) endpoint.

RESPONSE

CBR measurements are a useful method for assessing narcosis as a toxicity endpoint. They are, however, unsuitable for use when the substance has a specific receptor mechanism of action such as for pesticides in target organisms. Narcosis is normally observed at much higher concentrations than receptor-mediated responses and is often observed in non-target toxicity. Many of the PAHs act as narcotic agents and it is recognized that additivity of CBRs is an appropriate method for assessing the likely acute effects of PAHs in aquatic organisms. PAHs have been shown in recent unpublished work to cause increases in oxidative stress in fish (Hodson 1998). This stress leads to a number of responses that are similar to those mediated by the AhR. Once these processes are better understood, this may be another useful way to assess toxic potential of PAHs. However, carcinogenic potential is not well assessed using CBR. Many of the chlorinated pesticides (including some found at the site) are known to be toxic to arthropods and fish through receptor-mediated processes. Thus, these may have effects on arthropods and fish at body concentrations well below their CBR. An additive approach using narcosis to assess the chlorinated pesticides may not be appropriate, however, it should be applicable to the PCBs and similar substances.

B	Is the EPA 2 WQCTL approach (i.e., multiplying the Water Quality Criteria Chronic Value by the Bioconcentration factor) appropriate for determining ecological effects levels of the contaminants for which they were developed? Specifically, are the appropriate BCFs used (for fish, bivalves, etc)? (Please see Region2/CENAN joint evaluation memorandum, Appendix for Table 1, Page A-1)

RESPONSE

The EPA WQCTL approach for determining ecological effects levels was judged to be appropriate as were the BCFs used in these calculations.

C	BCFs reported for fish were used in the calculations of WQCTLs for organics; is this derived level appropriate for setting benthic tissue ecological effects levels? If the fish tissue levels are used, should adjustments be made to the derived levels to reflect the higher lipid contents of the benthic organisms used in the testing program?

RESPONSE

If WQCTLs based on BCF values measured in one organism with a very different lipid content than another, this may lead to incorrect estimation of tissue concentrations. Lipid normalization has been recommended (Connell 1990; Hebert and A 1995) in a number of situations and, in the experience of this reviewer, can significantly change interpretations. Lipid normalization should be used.

D	Are the WQCTLs calculated for metals using bivalve BCFs appropriate for setting levels for polychaetes or vice versa?

RESPONSE

Although this reviewer is not very familiar with metal toxicology, it is known that efficiency of metal uptake in molluscs can vary with food availability and is different from that in many other organisms (because of the intracellular digestive process in the hepatopancreas). Thus molluscs would be more efficient at taking up particulate metals (as particles or attached to particles) from the water-column. The application of BCFs for metals from bivalves to polychaetes is judged to be inappropriately conservative while the reverse is judged to underestimate potential for exposure potential in clams.

E	Are the uncertainty factors applied while deriving ecological effects levels for PAH contaminants appropriate? Does this adequately address the uncertainty around the derived values? Can uncertainty be accounted for using these order of magnitude adjustments? Should they be applied elsewhere to the other risk-based values?

RESPONSE

Uncertainty factors are used to account for unquantified uncertainty and, as such cannot be judged against the true uncertainty (until this is known). Order of magnitude factors are frequently used for animal-animal extrapolation and to account for unknown variability in population responses. They are no substitutes for a knowledge of variability and uncertainty, however, "arbitrary" uncertainty factors of this magnitude have been successfully used in the past and their continued use in the face of insufficient knowledge is judged appropriate.

	F	Are the risk values suitable for predicting the significant undesirable effects due to bioaccumulation; are there better alternatives for ecological nonspecific risk?

RESPONSE

The hazard quotients used in this assessment are judged appropriate. See the discussion of risk and hazard above.

	G	If you believe that these values are over- or under- conservative, what do you believe to be an appropriate way to improve them.)

RESPONSE

NA

CALCULATIONS

9		Should total PCBs continue to be estimated by doubling the total of 22 congeners or should it be quantified directly using another measure of quantification? What method is most appropriate for sediments in the NY/NJ Harbor area? (Please see Reference No. 60, Table 4-4B)

RESPONSE

This reviewer is not familiar with recent advances in the analysis of PCBs, however, the doubling to account for unquantified congeners seems a reasonable approach as it is based on historical experience.

10		Currently, 28-day tissue concentrations of certain organic contaminants are adjusted by some multiplier to estimate the concentrations of those compounds had the exposure been of sufficient duration to allow attainment of steady state levels. (Please see Reference Nos.5 and 46) Are these adjustments appropriate? Should steady state corrections be applied to any other of the listed contaminants? Are there other compounds for which we test that are not expected to approach steady state within the 28-day period?

RESPONSE

The use of a multiplier to estimate the equilibrium concentrations of those compounds that have not reached steady state levels in 28-d exposures is judged to be reasonable, based on observations and experiences with experimental studies where long-term body-burdens have been measured (Lee et al.

1994; Pruell et al. 1993). As a general rule of thumb, correction factors should be applied where log K_{OW} is greater than 4 and half-life for depuration from the tissue is more than 9 days.

11 Is the calculation and use of BaP toxicity equivalence an appropriate way to estimate the potential carcinogenicity of PAHs? (Please see Region2/CENAN joint evaluation memorandum, Appendix for Table 1, Section C.)

RESPONSE

BaP TEs are judged to be an appropriate method for estimating the carcinogenicity of PAHs. PAHs usually require metabolic activation as they are pro-carcinogens. With high exposures to mixtures of PAHs, metabolism may be reduced by substrate overload, thus lowering the carcinogenic risk. As exposures reported in this assessment are generally low, this is unlikely to occur, however, the qualifier discussed above in relation to extrapolation and repair mechanisms needs to be considered. The use of BaP TEQs is judged to be somewhat conservative.

12 Similar to PCBs, only a subset of those PAHS present in New York Harbor are measured for testing evaluation. How should the remainder be considered?

RESPONSE

In this reviewers experience, (with PAHs in creosote) the concentration of the 15 EPA priority PAHs follows the toxicity of the balance of the components of the mixture although, prior to weathering, the complete mixture is usually more toxic than would be predicted from the 15 priority PAHs. Given the age of the sediments in the site being assessed, the 15 priority PAHs are judged appropriate for estimating toxicity.

13 Is the assumption of a trophic transfer coefficient of one appropriate for use in evaluating the potential for human health and ecological impacts associated with metals in Remediation Material? Are the trophic transfer factors calculated for organic compounds correct? (Please see Region2/CENAN joint evaluation memorandum, Appendix for Table 1, Attachment C.)

RESPONSE

The trophic transfer factors used in evaluation of human and ecotoxicological health in this assessment are judged to be appropriate.

14	Is the assumption of a fish consumption rate of 6.5 g/day appropriate for use in evaluating the potential for human health impacts associated with metals in Remediation Material? (Please see Region2/CENAN joint evaluation memorandum, Appendix for Table 1, Page A-5) Would it be appropriate that the evaluation focus on a higher consumption population?

RESPONSE

Given the low likelihood that fish or shellfish directly from the site will be eaten by any particular individual on a consistent basis (no local subsistence fishery), this assumption of an average consumption of fish of 6.5 g/day is judged to be appropriate and probably conservative.

GENERAL

15	Is it plausible to replace any other risk assessment assumptions with assumptions specific to the HARS site? (Please see Region2/CENAN joint evaluation memorandum, Appendix for Table 1, Attachment C and Reference Nos. 88) Is it appropriate to consider the HARS intended use to be factored into an evaluation of effects at the community or population level?

RESPONSE

This reviewer believes that the assumptions used in this assessment are reasonable and consistent with other assessments of similar situations. Most of the criteria used in the assessment are aimed at individuals or individual populations. They are therefore judged to be sufficiently conservative to be protective of population and community responses.

16	Is use of the Squibb et al. (1991) report appropriate for identifying the contaminants of concern? Are there contaminants which should be added to or deleted from the list of contaminants for which we presently test? Please see Reference No. 51)

RESPONSE

The Toxics Characterization Report (Squibb et al. 1991) is judged to be appropriate for the identification of potential contaminants. Analyses of some compounds such as the minor metabolites of

DDT and some of the other pesticides is judged to be less necessary as they are less toxic, however, they are usually analyzed along with other analytes and the information would be available anyway.

17 Should risks from synergistic effects, from exposure to multiple contaminants, be evaluated using results from tissue analyses? If so, how? If not, why not?

RESPONSE

Additivity seems to be the rule where stressors are present at concentrations below their individual physiologically active concentrations. Toxic units are commonly used to assess such mixtures. The most appropriate uses of the toxic unit approaches are when the stressors are known to act additively. When the stressors are known to act independently, the hazard rate approach is more suitable. When the stressors are known to act synergistically, by potentiation, or by antagonism the use of multivariate procedures is more appropriate, however the data requirements may be large and empirical experimental techniques may be more appropriate. Pharmacologically based toxicodynamic models may be applicable in some instances where sufficient data are available (Kooijman and Bedaux 1996).

Although synergism and potentiation of substance-mediated responses are perceived to be a major concern in the assessment of many interactions, the likelihood of these occurring in the case of mixtures of substances in the environment is not as great as might be expected and neither is the degree of interaction. For example, Alabaster and Lloyd showed that the majority of toxic interactions between components of effluents were less than additive and that the likelihood of observing synergistic ratios greater than 8 was small (Alabaster and Lloyd 1980) Könemann and Pieters report that, in several studies on the toxicity of mixtures of substances where the individual components were present at specific fractions of a standardized response (e.g., LC50), the mixtures were never less toxic than the most toxic component and potentiation was not observed Könemann and Pieters 1996. Under the conditions of exposures to low concentrations that are described in this assessment, the most appropriate approach is to assume additivity. Synergism is judged to be unlikely to occur.

18 Is test tissue concentration exceeding reference tissue concentration by less than 10X a meaningful evaluative criterion? (Please see page 9 of the Region2/CENAN joint evaluation memorandum)?

RESPONSE

No, the choice of the reference can confound the results (see discussion on question 3 above).

19 Are the studies from which background tissue concentrations were calculated weighted appropriately? If not, what method is recommended? Is the use of the mean the most appropriate measurement of central tendency? If not, what measure should be used? (Please see Reference No. 98) Are the assumption, presented on page 14 pertaining to comparisons of bioaccumulation in test tissue to tissue concentrations in organisms from the vicinity of the remediation site, valid for evaluating undesirable effects?

RESPONSE

Where sufficient data are available, a distribution, rather than a mean should be used. This would allow probabilistic risk assessment techniques to be used. Where the underlying distribution of the data is known, a statistical measure of central tendency can be used (e.g., geometric mean of log-normally distributed data). However, the use of the central tendency in the absence of knowledge of the range or variance is counterintuitive - we should be more interested in the upper centiles of exposure and the lower centiles of sensitivity. Where the data sets are small and the underlying distribution is not know, the arithmetic mean is appropriately conservative. Where contributions to the whole are being calculated, only the arithmetic mean should be used. A recent paper by Parkhurst discuses this in more detail (Parkhurst 1998).

20 Can baseline tissue concentrations, from appropriate benthic organisms resident to the HARS, be used as standards to determine suitability for Remediation Material as defined above?

RESPONSE

Yes, with the qualifier on lipid normalization noted above.

5 REFERENCES

Alabaster JS, Lloyd R. 1980. Water quality criteria for freshwater fish. London: Butterworth. 200 p.

Connell DW. 1990. Bioaccumulation of Xeobiotic Compounds. Boca Raton, FL: CRC Press.

Hebert CE, A KK. 1995. To normalize or not to normalize? Fat is the question. Environmental Toxicology and Chemistry 14:801-807.

Hodson PV. 1998. Personal communication. .

Klaine SJ, Cobb GP, Dickerson RL, Dixon KR, Kendall RJ, Smith EE, Solomon KR. 1996. An ecological risk assessment for the use of the biocide, dibromonitrilopropionamide (DBNPA) in industrial cooling systems. Environmental Toxicology and Chemistry 15:21-30.

Könemann WH, Pieters MN. 1996. Confusion of concepts in mixture toxicology. Food and Chemical Toxicology 34:1025-1031.

Kooijman SALM, Bedaux JJM. 1996. The analysis of aquatic toxicity data. Amsterdam: VU University Press. 149 p.

Lee HI, Lincoff A, Boese BL, Cole FA, Ferraro SP, Lamberson JO, Ozretich RJ, Randall RC, Rukavina KR, Schults DW et al. 1994. Ecological risk assessment of marine sediments at the Unied Heckathorn superfund site. Narragansett, RI: US EPA ERL. Report nr EPA-600/X-94/029.

Parkhurst BR, Warren-Hicks W, Etchison T, Butcher JB, Cardwell RD, Volison J. 1995. Methodology for Aquatic Ecological Risk Assessment. Final Report prepared for the Water Environment Research Foundation. Alexandria, VA: Water Environment Research Foundation. Report nr RP91-AER.

Parkhurst D. 1998. Arithmetic versus geometric means for environmental concentration data. Environmental Science and Technology(Feb 1 1998):92A-98A.

Parrott JL, Hodson PV, Servos MR, Huestis SL, Dixon DG. 1995. Relative potencies of polychlorinated dibenzo-p-dioxins and dibenzofurans for inducing mixed function oxygenase activity in rainbow trout. Environmental Toxicology and Chemistry 14:1041-1050.

538 Pruell RJ, Rubinstein NI, Taplin BK, LiVolsi JA, Bowden RD. 1993. Accumulation of polychlorinated
539 organic contaminants from sediments by three benthic marine species. Archives of
540 Environmental Contamination and Toxicology 24:290-297.
541 Solomon KR, Baker DB, Richards P, Dixon KR, Klaine SJ, La Point TW, Kendall RJ, Giddings JM,
542 Giesy JP, Hall LWJ et al. 1996. Ecological risk assessment of atrazine in North American
543 surface waters. Environmental Toxicology and Chemistry 15:31-76.
544 Squibb KS, O'Conner JM, Kneip TJ. 1991. New York/New Jersey Harbour Estuary Program
545 Module 3.1: Toxics Characterization Report. Tuxedo, NY: Institute of Environmental Medicine,
546 New York University Medical Centre.
547 Stephan CE, Mount DI, Hansen DJ, Gentile JH, Chapman GA, Brungs WA. 1985. Guidelines for
548 deriving numerical national water quality criteria for the protection of aquatic organisms and
549 their uses. Duluth MN: US EPA ORD ERL. Report nr PB 85-227049.
550 USEPA. 1991. Evaluation of dredged material proposed for ocean disposal. Washington, DC: US
551 Environmental Protection Agency. Report nr EPA-503/8-91/001.