N00102.AR.001913
NSY PORTSMOUTH
5090.3a

LETTER AND COMMENTS REGARDING REVISED DRAFT FINAL ESTUARINE ECOLOGICAL RISK ASSESSMENT NSY PORTSMOUTH ME
8/18/1997

RECEIVED
AUG 21 1997

August 18, 1997

Carolyn Lepage
Lepage Environmental Services, Inc.
P.O. Box 1195
Auburn, Maine 04211-1195

Dear Ms. Lepage:

I have reviewed the report entitled Revised Draft Final Estuarine Ecological Risk Assessment for the Portsmouth Naval Shipyard Kittery, Maine April 7, 1997. Comments on the document are provided in the attachment to this letter.

The comments address the questions from The Seacoast Anti-Pollution League, April 12, 1997, about the document and the site and a critical review of the scientific basis of the document.

Although several points have been identified that require additional work, overall the document and the research done contribute much to the understanding of the ecological and environmental risks from the Shipyard. Those scientists involved are to be complemented on difficult task which has come far.

I realize that the comments are fairly detailed and at times complex. I would be pleased to provide any clarification which may be needed to understand them.

I appreciate the opportunity ti assist in this interesting and important project.

Sincerely

David R. Brown Sc.D.
Westport Connecticut 06880

COMMENTS ON THE ECOLOGICAL RISK ASSESSMENT FOR PORSMOUTH NAVAL SHIPYARD AUGUST 17, 1997

The following comments on the Revised Draft Final Estuarine Ecological Risk Assessment for Portsmouth Naval Shipyard, Kettery, Maine, Volumes 1 and 2 were prepared by David R. Brown, ScD. on behalf of the Seacoast Anti-Pollution League (SAPL) Portsmouth, New Hampshire, as the subcontract to Lepage Environmental Services, Inc., Auburn, Maine.

Introduction

The review covers the complete report with detailed consideration of sections on the ecological/toxicological areas. Comments and recommendations are provided with respect to the methodology used in the report and interpretation of the risk.

SAPL formulated eleven (11) questions about the report or the findings that were of concern. These questions are addressed in Part I of the review. Part II is an overall evaluation of the Ecological Risk Assessment and the potential significance to further clean-up actions. Part III is specific comments and recommendations for each section of the report.

General

The Assessment follows the EPA Ecological Risk assessment format for superfund sites. The ERA describes an extensive series of investigations into the ecological conditions at the site and estuary that provide valuable information about potential environmental impacts. The document is complex. It contains extensive technical discussion and complicated, data-rich tables. It is difficult for the lay reader to understand and to evaluate information in the report. The report needs a clear explanation of the findings and the implications to future work. This should be provided both in Section 1, Executive Summary, and in Section 8, Summary, Conclusions, and Recommendations.

PART I.
SAPL Questions

An objective of the ERA is risk communication to the public. A SAPL committee developed a list of questions that illustrate some of the public concerns. These questions concern the methods of analysis, accuracy of the data, completeness of the report, and overall risk to health and the environment. The following addresses the eleven (11) specific questions from SAPL as covered in the ERA.

1. Where are the "hot spot" areas of highest contamination and highest impact on the ecological systems?

Based on Table 8-1 (page 8-25), Clark Cove, Sullivan Point, Dry Dock, Back Channel and Jamaica Cove are the areas of highest ecological concern. Points of contamination within these areas that could be designated as "hot spots" of chemical contamination are not apparent from discussions in either the conclusions or risk characterization sections. Review of data in the appendix suggests that there are areas that could have higher contamination.

correlation not cause & effect

The highest impacts on ecological systems are in the areas near the seeps at Clark Cove, Sullivan Point and Jamaica Cove. It was found that a) mussels accumulated lead and phenanthrene (PHEN); b) metals affected eelgrass root growth; c) sea urchin reproductivity is altered by lead and d) DDT toxicity is seen in amphipods. suspected The areas closest to the SWMU appear at greatest risk. Risks in areas further from the shipyard are not ruled out, but are less likely. When there is ecological impact, it appears to be to a specific fauna or flora rather than whole ecosystems. See comments in Part II and III for further discussion of this point.

2. Does the ecological risk assessment use averaging such that the hot spots are not still evident?

Yes, in Sections 7 Risk Characterization and in Section 8 Summary Conclusions and Recommendations, the statistical analysis includes averaging that would tend to obscure evidence of a hot spot. This is true both for areas of high contamination and for effects on specific parts of an ecological system.

Examples can be seen in Tables X-2 (page X-10) in which the mean, standard deviation, minimum and maximum values for some contaminants are shown. The data indicates an uneven and skewed distribution (see lead at Jamaica Cove, Back Channel, and Dry Dock; zinc at Clarks Cove and in table X-3; cadmium and arsenic at Clarks Cove). Even though the analysis in the risk characterization transformed the data to a "log" scale for statistical testing, the tests for statistical significance lacked the power needed to show differences when they exist. Figures 7-1 and following figures show that when using weight-of-evidence analysis, potential effects are reduced in weight because of averaging with other effects assigned lower weights (Figure 7-1, 7-5, 7-6, 7-10). The same is true for the analysis of potential exposures. The source of the skewed data problem appears to be related to the occurrence of seeps near an Area of Concern but may also occur elsewhere for different reasons.

The question of "hot spots" should have been considered in the Risk Characterization narrative and in the limitations and uncertainty discussions. They are not. "Hot Spots" should also be considered in the Risk Management and decision making sections

3. Has the data been gathered at enough distances from the shipyard to assess damage to areas such as Spruce Creek, Fort Foster or the Portsmouth side of the river?

Data is available (Fig. 4-12) for areas distant from the shipyard. Models of distribution have been run that show the overall distribution in the estuary, Fig 5-1. The statistical analysis does not specifically consider each of these regions. The data available for Spruce Creek suggests a possible impact from the JC and CC landfills and the seeps. The rate of tidal flow through the estuary may be a reason for less concern on the Portsmouth side of the river. It is not clear from the report whether sufficient data has been collected to assess the damage there.

4. Is data sufficient to give an accurate assessment of the risk with respect of sample numbers of lobster, shellfish and test wells?

An accurate assessment requires both strong data and strong qualitative analysis of the data. For the most part, the data collected is abundant and gives a relatively complete picture of the contamination. There are some examples of small sample sizes, such as seep evaluation, where as few as two samples are used to support conclusions. Overall, the sample sizes tend to be small, but corroborative samples in other media such as mussels, sediment and pore water add strength to the analysis and conclusions.

The analysis of the data in the assessment is a problem. The assessment, as presented, does not give a complete picture of the potential areas of concern. Data that fall outside of the statistical ranges are not treated as "hot spots." This concern could be addressed by additional targeted sampling, the design of which is based on the evaluation of present data and by a monitoring program.

5. What effects have the yard pollutants had on Spruce Creek?

Spruce Creek appears to have elevated levels of contaminants in the sediment which, based on the model and the sediment data, appears to be linked to Jamaica Cove and Clark Cove seeps. There are increased levels of metals in mussels; but the impact on the mussel population in Spruce Creek is not clear from the report. Other impacts may be present in Spruce Creek but not related to the Shipyard. Comparison of surface and deep sediment samples suggests that, for some contaminants of concern, surface sediment is as high or higher than deep sediments in both Spruce Creek and areas with seeps (Tables 5-4 to 5-7). This may indicate that there are still releases into the system.

6. Is there a direct connection between a source at the shipyard and a contaminated site off shore?

The report does not provide sufficient information to address this question. Sampling of lobster and winter flounder do not give an indication of an effect. It may be possible to identify off shore locations for further study from the modeling results. Some evidence of low levels of contaminants of concern are found in lobster and some flounder but sources of exposure are unknown.

7. Has there been an investigation and are there reports of radioactive contamination on Seavey Island?
None are reported in the ecological risk assessment but there may be information from other studies.

8. Have potential problems been ignored because of an existing condition?
The report does not exclude any areas from the analysis because of preexisting conditions such as closure of shellfish beds or fish advisories.

9. Is the total accumulated risk to environment from all sources in addition to the shipyard considered and made clear in the report?
The impact on the ecological system from other sources of contamination in Great Bay and Piscataqua river is not addressed in the report. Therefore, the relative contribution from the shipyard is unknown.

The weight-of-evidence (WOE) tried to consider total accumulated risk from the shipyard only. It is a difficult task and the risk characterization only had limited success towards this goal. One problem is the uncertainty in many of the data sets and thus in the conclusions. To obtain total accumulated risk in a quantitative sense involves adding all of the risks together. There are two problems with that approach. First, the risk levels identified have factors for uncertainty in them, e.g. safety factors of 10 and/or more. When risks are added together the safety factors are also added. The uncertainty that is added when summing the risks contributes as much or more than the actual risk to the total. The numbers obtained are not meaningful. The other difficulty is that the different types of risks are not necessarily additive. For example, risk to lobsters and mussels are different from risks to eelgrass or salt marsh even though they interact in the environment.

However, it is possible to address the issue of accumulated risk in a qualitative manner. A discussion section could be added to the report that focused on the interrelationships and the relative concerns from all risks. There is some attempt to do this discussion in section 8, but it should be more helpful in separate section.

10. Is it safe to swim in Back Channel?
This question is more properly addressed by the local Health Department. There are contaminants in the water and sediment of the Back Channel, although they are not as high as other areas. Some of these contaminants at high concentrations could induce skin rashes and possibly other acute effects. Since the testing indicated that there are spots of higher contamination due to local conditions in Back Channel, a health department evaluation before swimming in Back Channel may be necessary.

11. Is it safe to fish near the shipyard?
Again, this is a question for the Maine Health and/or Environmental Departments. Fish sampling data is needed to answer this question. Maine and New Hampshire both issue fish consumption advisories and FDA monitors commercial fish to assure safety. It

should be noted that the rates of accumulation of toxic chemicals in fish are relatively specific for each species and habitat. While winter flounder, lobster and mussel did accumulate some metals and organics in this study, the levels are low and within levels considered to be safe.

PART II

Overview and General Comments

1. The Ecological Risk Assessment is logically organized and follows the EPA guidelines for ecological risk assessment at superfund sites. The risk characterization does not bring the issues together in a way that the inter-relationships between risks can be evaluated. The general prioritization of the risks cited in the conclusions is biased down towards reduced risks. Assignment of lower risks as presented in the conclusions limits the risk management options.

2. The statistical treatment of the data needs to be reconsidered. Detailed examination of the data in the report and the appendix indicates serious limitations in the analysis. Tables in both the report and the appendix need improved legends and footnoting. Parametric statistical approaches are being applied to non-normally distributed data. Inspection of the raw data in Appendix 2 reveals that the data is not only skewed but in some cases it appears to be bi-modal or multi-modal. Given the small sample sizes, even a log-normal transformation may not be sufficient to support the statistical treatments used. It should be possible to analyze some of the information with non-parametric statistics. As presented, the statistical probabilities infer an accuracy and precision that is not correct.

3. There is a very rich data set available and the sampling strategies are robust. Much of the strength of the sampling strategy is the inter-relationship between the measures such as pore water, sediment and mussel uptake of metals. The relationship is further enhanced by the ecological effects measures. However, the relationships between parameters are not discussed. Correlation coefficients were calculated and shown in the appendix, but the findings are not incorporated in the weight-of-evidence analysis.

4. The attempt to evaluate the confidence in the data in mathematical or statistical terms goes too far. The discussions in Appendix IX are more informative. But a major concern here is the handwritten entries on the tables in Appendix IX which suggest that these are draft documents and it is not known which are the correct entries. When the terms high or low confidence are assigned to a conclusion, the reader is presented with a paradox; should he/she reject the finding or ask for more data or assume that there is no risk?

5. The toxicology presented and the derivation of the effect levels such as LC50 and its use in the probability discussion needs to be reconsidered. The toxicity discussions are insufficient basis for judgment of risk. The slopes of dose response curves are needed as a minimum. Also, there needs to be uncertainty factors incorporated before the data are used in risk assessment. Most important, when there is sparse toxicity data it is very difficult to verify that the sampling strategies are correct, or that plausible effects are being observed in the ecological evaluation. A more complete toxicology profile is needed, such as those developed in the ATSDR toxicity profiles.

6. The conclusions in section 8 indicate issues of concern. But the prioritization given is biased toward the low end. They are as follows:

- The risk assessment concludes that there are intermediate, low or negligible risks to receptors in certain ecological systems. It further concludes that no area has high risks.
- It is concluded that the risks emanate from the sediments and the surface water with chemical contamination.
- Links were ~~shown between~~ certain SWMU and ecological systems.
- Chemicals involved include both heavy metals and organics.
- It concludes that the risks are near Seavey Island and are not a concern for the greater estuarine system.
- With respect to ecological findings, risks were identified for flounder in the pelagic community, mussels in the benthic community, as well as eelgrass, salt marsh and avian species.
- When ambient conditions are considered, risks of exposure and bioaccumulation for Portsmouth estuary are reported to be comparable to some risks found in reference regions.

If these findings are to form the basis of the remedial activities at the site it is important to examine the scientific basis for them. The report needs to be strengthened to show how each of these conclusions flow from the data collected. It needs to also show how the data supports the rejections of risks that are not included in the report. Most importantly, the prioritization in Section 8 needs to be consistent with field and sampling observations.

Part III
ERA Specific comments on sections of report

The following are specific comments, questions and recommendations for each section of the report and appendices. EPA and the states of Maine and New Hampshire provided comments to an earlier draft of this report (Appendix I). Those comments and recommendations are not repeated in this peer review. However, in several instances, it is noted that certain prior comments have not been addressed in the present draft. The review considers method of analysis, data quality, and interpretation of the findings.

Section 1. Executive Summary

1. **General:** The objectives of the ERA are to determine the adverse effects of shipyard waste on the Ecology of the Piscataqua river and the estuary. The executive summary is meant to inform the less technical reader and should give a broad but concise overview of the work. This summary considers the approach to the research, the methodology and the conclusions. Most material is presented in outline form. The summary is not helpful in providing information to the general public or to the media. A risk manager would have to review the entire report in detail to understand the basis for most of the conclusions and to develop a risk management strategy.

2. **Questions and comments:**

a. Often the Executive summary is the only section read in reports of this length. Does the this summary contain enough information to address the 5 W's (who, what, when, where and why)?

b. The approach to the research is described in detail but the conclusions are given in abbreviated form without addressing their meaning. The limitations in the ERA are not presented at all in the Executive Summary although they are very important to the risk management decisions.

c. The section should also discuss those risks that are ruled out by the assessment.

3. Recommendations:

Expand this section to give a more complete picture of what was done, what was found in the sampling and what the conclusions mean. This is excellent work that would be enhanced by a more complete analysis of the data and discussion of the findings. This report should contain a section that ties the work together or a separate paper should be prepared to place the entire effort in perspective.

Section 2 Introduction

1.General: In addition to the assessment of the risk and development of information to support risk management,, an objective of this report is to support the communication of Shipyard-associated ecological risks to the public. The ERA scope includes problem formulation, exposure and ecological assessment and characterization of the risks. The objectives of the ERA are nicely set forth in this section of the report.

2. Questions and Comments:

a) Is the material in the appendices considered to be an integral part of this report? Will the public be given both Parts I and II when they request to report?

b) A risk communication discussion would be helpful. This should include the expected questions from the public and answers plus a media interaction plan, which could interpret the findings for the press.

3. Recommendations:
It would be helpful to state the questions are that the report is designed to address. If the report is to be a descriptive document it will be different from a document that addresses scientific research questions. For example, the type of information needed to determine whether an exposure is higher than established regulatory standards is different from that needed to develop a hypothesis to be tested in experimental studies. In some cases this report contains experimental research while in others the report simply presents analytical findings without conclusions or interpretation.

A clear statement of what the report is doing and is not doing would reduce confusion and perhaps prevent misuse of the findings.

Section 3 Site Description

1.General: This is a physical discussion of the Shipyard and the Great Bay Piscataqua River estuarine system, the Solid waste management units, potentially exposed habitats, and potentially exposed species. It is very helpful and raises questions that should be addressed in the report. Other sources of contamination that may be present in the estuary are not presented. The absence of this information leads to confusion in understanding the certain findings in the report. For example, The sewage treatment plants and manufacturing are not considered as sources.

2. Questions and Comments:

a) Table 3-1 shows SWMUs being investigated. Are these the only units or are there others which are not included? If so why were they omitted?

b) The shipyard serviced nuclear vessels. Was there radioactive material disposed of at the site or held in storage at the site which could have reached the estuary?

c) Which of the Hazardous materials listed in table 3-1 are now shown to be present at the SWMU?

d) Are the 55 gallon drums in which waste and solvent were placed in the Jamaica island landfill still in tact or have they begun to leak due to deterioration?

e) There appears to be a discrepancy between the findings at JC landfill by the two different contractors. Is there evidence of material being released from the land fill from the seep analysis? Is there sufficient data to determine the presence or absence of releases? At what stages of the tide are they released?

f) SWMU 5 and 6 sampling of biota and soil has shown metals present in both. Does this mean that materials are being released from the SWMU and taken up by biota? Is there sufficient information to address this concern?

g) Seven SWMUs were identified with no follow up. Were these dropped from additional consideration in the ERA.

h) The intertidal mudflats and the and the rock intertidal habitat are not specifically discussed in the ERA risk characterization. Are these habitats considered under a different name

i) Are flounder studies sufficient to evaluate all of the fish species in table 3-8? This table needs a more complete legend.

j) Are the ponds on the island salt or fresh water? Is there a tidal connection?

3. Recommendations:
A short summary of the most important points in this section would assist the reader in understanding the important issues in the site description. For example, a statement that the assessment will focus on certain SWMU and ecological systems for the following reasons would help set the stage for the following sections. As it is currently presented, the reader has to infer what is important from reading the chapter.
As written, this a very informative section and helps to explain the sampling design.

Section 4. Selection of Assessment Endpoints, Contaminants of concern and measures of exposure and effects

1. General: Table 4-5, figures 4-2, 4-12 and 4-22 form the basis for and understanding of this section of the report. Prior comments raise several important issues relative to this section (Appendix I). Formulation of the risk assessment requires the selection of correct end points, measures of exposure and effects, weight-of-evidence procedures, a conceptual model, and an exposure pathway analysis. This is key to the success of the project. It is not clear from the discussion how the selection was done. The process is probably valid but several points need to be clarified, especially with respect to the weight-of evidence analysis.

2. Questions and comments:

a) 1. WOE emphasizes the uncertainty of the assessment data. Wouldn't it be more logical to base the risk assessment on what is known and plausible rather than on the uncertainty? The discussion in appendix IX is much more informative than the tables on weight-of-evidence (WOE).

b) The Conceptual model discussion is helpful, especially the last part of paragraph 3 page 4-5 in which transport and movement through the biota and water system is shown.

c) Table 4-5 is very important but it is difficult to understand. Does the number of bullets indicate the number of statistically significant tests in which mussels accumulated COC ?

d) It is difficult to determine the basis for the ER-L measures and how they are used. (See footnote, Toxicity thresholds on page 4-41.) Has the selection of these thresholds been reviewed by a toxicologist? The data in table VII-1 to VII-12 is not sufficient to support the ER-L measures. It must be clear how they are derived. How can a laboratory derived LC 50 be used to screen an *in situ* exposure?

e) It is stated that the ER-L tests are conservative because the binding to sediment was not used in the laboratory. Will binding will increase toxicity if the sediment is ingested?

f) Inspection of the raw data in table III-1 and following, shows that the field samples are sometimes highly variable especially in the reference areas. The statistical treatment of these data must reviewed. It is possible that use on a non-parametric statistic will give different results.

g) In general, the COC found are likely correct for all sites together, but the COC assigned to each AOC are less certain. What statistical criteria was used to verify that a data set could be used for analysis of COC at each AOC? (e.g. the F test for equal variance or normality of the distribution?)

h) The calculation and use of the HQ is not clear. How are these values obtained and how were they used? Is there any instance where a HQ or HI was not calculated due to data limitations or sample size or variability?

3. Recommendation:

Although there are questions about this section, the findings appear to be about right. This is probably the result of the use of multiple screening approaches. Because these findings form the basis for the entire analysis in the remainder of the ERA, more verification is needed. Problems in the risk characterization can be traced back to decisions made in this section. Although it is a difficult task the critical decision points need to be identified and defended in the discussion. There is excellent work here but the interpretation of the findings is incomplete. Application of statistics to inappropriate data should not be a basis for excluding information needed for complete understanding of the ecological risk.

Section 5 Exposure Assessment:

1. General: The analysis in this section shows differential movement of contaminants of concern (table5-3). The concentrations listed in the environmental media table 5-4 are central to the interpretation and understanding of the risk at the sites. The section needs to discuss these findings in both an analytical and a descriptive manner. If one simply looks at the data in the tables some interesting questions begin to emerge. First, the sample sizes are small, maximum values are far in excess of the average, and the STD is approaching the average. The data is not normal and may not be log-normal. More thought needs to given to the statistical analysis. Wouldn't a non-parametric test have been useful for non-normally distributed data? Much more information can be gleaned from this section.

It appears important that the surface sediment samples of some COCs are nearly the same as the core samples. Would this be true if the releases to the environment have been controlled for some time?

2. Questions and Comments:

a) See prior comments from EPA and State of Maine and note the questions from SAPL at the beginning of this review.

b) What solid evidence is there to evaluate the penetration of contaminants into Spruce Creek? The modeling suggests that material can enter the creek.
c) Note the data in figure 5-12 with respect to the marsh. Does this mean that the marsh is an area of concern?
d) Do the figures 5-8 through 5-24 support the conclusions in the risk characterization section 8 of low or intermediate risk levels?

3. Recommendations:
The data in this section is relatively strong. Although it is difficult to analyze mathematically or statistically, it provides a nice picture of what is going on in the system. While it is very important to identify data gaps and limitations as was done in the report, it equally important to interpret the data available. This section should be reviewed from the perspective of the questions being asked in the ERA and those needed for risk management decisions.

Section 6: Ecological Effects

1. **General:** The inter-relationships between components of the ecological studies and concentration of contaminants needs to be shown. This chapter uses three ways to measure ecological impacts : 1) Published effects of contaminants identified at SWMUs on each type of receptor. (The concentration of contaminants believed to adversely affect parts of the ecosystem), 2) Evaluation of ecological changes at each site, and 3) Criteria and standards for toxic changes from the literature.

   These measures are compared to exposure levels determined in Section 5. Ecological effects are found in the following: sea urchins, deployed mussels, winter founder (BAP accumulation and spleen damage), and lobsters (PAHs). In indigenous mussels elevated levels of contaminants but no physiological effects are found. (However, laboratory studies suggest a steady state in mussel uptake with sediments and water levels.) Rockweed accumulated lead and copper. Other partial findings are:

- Eelgrass rhizome length reduced by lead.
- Salt marsh trapping of contaminants.
- Avian Dietary hazard quotient, (a summary of hazard quotient for each chemical in the index) that suggests a negligible effect.
- Dose-response effects of lead in the benthic community in laboratory experiments.

The relationship between the various findings needs to be discussed.

2. Questions and comments:
a. Field studies and laboratory studies are used to evaluate the contaminants of concern, the effects and the exposures. This section needs to add a discussion of the inter-relationships between field and laboratory studies.
b. A discussion of the studies conducted with statements relative to the confidence in the data and study design as well as relevance for evaluation of ecological risk is needed.

In addition, the presentation of findings and conclusions need to be expanded. In most cases there is only a brief sentence of interpretation for the findings. In a few instances not even that. Referral to Appendix IX provides tabular conclusions but there is no discussion of the meaning of these findings either.

c. Although for Fig 6-6 the requirements to use the model are satisfied, the range of lead values used in the laboratory studies is too large. The differences in exposures range over two orders of magnitude. It is unlikely that ED20 and ED50 values obtained have any meaning in terms of the response to lead in the field. These benchmarks cannot be used for the prioritization of risk because the levels of lead in the laboratory are too different from those observed in the field studies.

d. Exposures are presented that appear to be important but the discussion of the data is not complete enough to allow interpretation of the actual significance of the findings.

e. The discussions in Appendix IX are informative. Can the author expand them to give some interpretation beyond the brief statement of conclusions.

f. The Duncan's multi-range test statistic is given in some of the tables for the field or laboratory studies. What does this mean? Do the authors believe that the findings are ecologically relevant?

g. A clear identification of the findings and conclusions needs to be provided for much of this data. The tables in Appendix IX give brief mention of the actual effects. But it is too brief to support conclusions.

h. The tables in the review draft have handwritten changes. Which interpretations are the correct ones? Why were they changed?

i. Responses at sites are compared using existing standards; WQL thresholds, ER-L and for sediments, 10 and 50th percentile responses are used (Table 6-12). This table and table 6-14 are complicated and confusing. Can the information be clarified?

Recommendations:
The findings need to be tied together in a discussion. Same data is based on laboratory studies that has little direct relationship to the question of ecological impact. The section requires a comprehensive discussion of the significance of the findings.

Section 7 Risk Characterization

1. General: Contaminants of concern and ecological changes are found in most of the areas and presumed to be associated with the site, yet the analysis in the risk characterization chapter determined that there is no problem at most areas of concern. The only ecological risks identified are characterized as negligible to intermediate. Careful examination of the methods used and the data used for the analysis raises serious issues with respect to these determinations.

The characterization depends heavily on the Weight-of-evidence framework to evaluate the effects of the exposures. Information was analyzed from:

- Exposure in environmental media
- Exposure levels relative to measured effects

· Accumulation of contaminants of concern in biota and media
· Ecological status of components in the areas of concern (AOC)

The weight-of-evidence procedure compared three different types of factors: measures of exposure, effects, and the assessment endpoints. Linkages between SWMU and effects at the AOC were determined and assigned levels of confidence. The approach used canceled out important high risks in one factor with lesser outcomes in other factors. It assumes all factors are of equal weight and did not consider abnormal data distribution.

2. **Comments:**

a. Weight-of-evidence attempted to incorporate: a) strength (plausibility) of the data, b) quality of the data and c) study design and to attribute a composite score. This score is used to designate the confidence in the risk discussion but appears to have heavily influenced the determination of important risks. In the Portsmouth risk assessment this approach significantly reduced both the number of risks identified and the seriousness of the risks found. See tables 4-7 to 4-10.

b. Conclusions from the risk characterization are based on statistical and other mathematical analyses that are inconsistent with the data. See tables 7-4 through 7-10. In these tables most risks are designated as negligible or low (22). Only 6 are at the intermediate level and none are designated at the high level. However, in Section 4, exposure and tissue accumulation of toxics are shown in water, sediment and mussels (Figures 4-22, 4-21, 4-19 4-12 and table 4-5). This accumulation is well over background or ambient levels in reference areas.

The explanation for the paradox between the conclusion and actual exposures is found with careful examination of the methodology used in the risk characterization. The following examples illustrate the problem.

Methodological Problem 1: Assignment of weighting factors to unequal measures.

Numerical weights are assigned to measures of effect, measures of exposure and endpoints. The numerical weights are next used to calculate the "centroids" found in figures 7-7 to 7-35. The location of the centroids is used to estimate the magnitude of risk from table 7-2.

These values are not linear and cannot be added together as is done in the figures. Doing so implies that a biological parameter such as specie density is directly comparable other parameters such as amphipod toxicity or pore water contaminant levels. If there is any relationship it is not likely to be linear. A linear relationship would be needed for the mathematical methodology used.

When risks are averaged, low risks cancel out the higher risks, even when there is no relationship between them. The higher risk which may in fact be the more important is discounted in the weight-of-evidence approach.

Methodological Problem 2: Assignment of weighting factors to unrelated measures.

Measures of effects, measures of exposures and end point weights are each assigned a numerical weight from 1 to 4. They are than treated as if they are comparable in the determination of the magnitude of risk. The weight assigned for a medium risk in end point category is not equivalent to the "low" designation for exposure categories of the "potential" for effects category. Thus a condition is created where a very important end point could be omitted from consideration because of poorly conducted measures of exposure or effects which have a lower confidence designation due to data quality.

Methodological Problem 3: Relationship of components within ecological systems.

Different observations within the endpoint, exposure and effects categories in figures 7-1 to 7-29 are treated as if they are equivalent. This is seen in Clark Cove (Figure 7-2). Seeps and mussel residue findings are assigned 'high' for exposure and 'medium' for end point, risks that are very important. But the overall epibenthic risk was assigned a 'low' magnitude of risk. There are other examples of the same type error in the figures. It is apparent that important indicators of problems are again being lost because of the analytical approach.

Methodological Problem 4: Assumption of equivalence in "cut off" values (benchmark, reference, etc.).

The estimation of risk from different media (water, sediment etc.) is altered by the evaluation of the quality of the data and the data selection in the weight-of-evidence. In tables 7-12 and 7-13 the magnitude of risk from exposure is related to the assessment endpoint numerically and those that exceed a cut-off value are assigned a level of risk. It is not clear how table 7-13 is used to weight the calculations. But again equivalence is assumed between different types of end points. Thus the accumulation of contaminants of concern by mussels is given a 'low magnitude of risk' for surface water although it is clearly shown to occur.

Methodological problem 5: Probability of exceeding benchmark.

The size of the sample data set affects the probability of exceeding a benchmark or other reference level. Review of the data in table 7-16 and multiple tables in appendix X show sample sizes that range from very small 2 or 3 to higher one of 30 to 50. Distributions and probability calculations appear to have been done without regard to sample size. For example see table X-2 and table 7-17 for Clark Cove where a sample size of two is used. When sample sizes are very low they have a strong influence on the variability of the data, and decrease the chance of finding a difference in a test of probability even when a difference is present.

Methodological Problem 6: Binding capacity for divalent metals.

It is pointed out that in risk characterization that the form of lead is an important factor in determining the toxicity. This is true for many divalent metals. *In vivo* toxicity is altered by AVS binding capacity and the test system whether Bulk or SEM. Toxicity is altered also by the form of the lead in the water or sediment. Lead carbonate, lead phosphate, etc. have different solubilities in the laboratory than lead nitrate. The actual chemical species of lead found in the sediment is not given in the tables. Further, trends are inferred from regression calculations which use data of different types (for example, VII-6 and elsewhere). It is not clear what the statistical regressions mean. When such data are used in comparison with the findings in the exposure study from Section 4, it is not possible to determine if the comparisons given are valid.

Methodological problem 7: Narratives.

The narratives 7.1 basically restate the findings in the tables and do not amplify the test. The narrative must clarify the potential short- comings of the statistical analysis other than mentioning the effect on uncertainties. The reader is left to rely on the mathematical analysis in the tables and figures for all conclusions. The narrative should place perspective on the findings and indicate where caution is needed when drawing of conclusions from the findings. A much better discussion is given in Appendix IX, but that discussion also needs to be enhanced. The data is too variable for the risk characterization to be based only on mathematical statistics

3. Recommendations:

Much of the data here is not definitive but the risks given in the report suggest limited ecological risk. The risk characterization and the narrative summaries should address plausible, if yet to be proven, concerns. Some of these can be found in the SAPL questions at the beginning of this review. Lack of definitive information in the fields of health and welfare is generally considered a reason for preventive and protective health and ecological actions.

One table in the report, table 8-2, presents significant findings for assessment and measurement endpoints and identifies research needed and essential monitoring studies. The table is excellent; it should be supported by the findings and discussion in the risk characterization section and form the basis for the section.

This section on risk characterization should be redone. Figures 4-22 and 6-1 suggest a way to integrate the discussion of risk.

Section 8 Conclusions and Recommendations

The findings from this section form the basis of the remedial activities at the site and thus it is important to fully examine the scientific basis for them.

1. **General**: While no major toxicological impacts or wide-spread environmental impacts are detected, indications of chronic risks are present. Overall findings are summarized in Table 8-1, page 8-25. The surface water and sediment at five (5) areas of concern (Clark Cove, Sullivan Point, Dry Dock, Back Channel and Jamaica Cove) have risks linked to metals or organics (PAHs, halogenated hydrocarbons) and to seeps for the higher risks. Compounds related to the pesticide DDT are found to increase risk at all areas of concern. Even though lead is present at all sites, the document determines that lead does not induce an ecological effect. Thus the conclusion of "overall no major toxicological impacts or wide-spread contamination were detected."

The ecological risk assessment identified areas of concern for intermediate risk, low risk, and negligible risk. No area is identified as having a high risk using the endpoints evaluated. Risks at certain areas of concern are linked to sediment and solid waste management units (SWMUs). Those chemicals that exceed the reference values presented in the risk assessment are designated as potential risk drivers. The pesticide DDT is present at all areas of concern and is a potential risk driver, but DDT is not linked to any SWMU.

Thus, the risk characterization concluded:

a. that there are intermediate, low or negligible risk to receptors in certain ecological systems and that there are no high ecological risks.
b. that the risks emanate from the sediments and the surface water with chemical contamination at some areas of concern.
c. that there are links between certain SWMU and ecological systems.
d. that the chemicals involved include heavy metals, organics, PAHs, and PCBs.
e. that the risks are near Seavey Island and no risks of concern to the greater estuarine system.
f. that with respect to ecological findings, risks were identified for flounders in the pelagic community, the benthic community (mussels), eel grass, salt marsh, and avian species.
g. that when ambient conditions are considered, risks of exposure and bioaccumulation for Portsmouth estuary are comparable to risks found in reference regions.

Careful analysis of the Risk Characterization process and supporting data indicate that more description of the conditions at the site are needed for risk management decisions.

2. **Questions and Comments**:
   a. It is not clear what the basis is for and the strength of the conclusions in table 8-1 on levels of risk. Findings presented in Appendix IX are not consistent with the conclusions from the weight of evidence evaluation. For example, juvenile lobster tissue residues and individual mussel tissue residue are elevated. How is that

finding in Appendix IX consistent with the conclusion of low or intermediate risk in Section 8?

b. There is a problem with the Hazard Index (HI) used which is based on highly variable data. Only HI values that exceed 1 or 10 were considered as positive evidence of risk, but no Hi exceeded 10. The HQ values were developed from receptor-specific concentrations in Appendix VII and modeled exposures in Appendix VIII. The effect on the variability of the data need to be considered in this analysis. Are averages used in this evaluation? Note earlier comment from EPA on this point.

c. Although no source can be found for DDT, the presence of DDT at several locations near the shipyard suggests sources at the site. How was DDT ruled out as a site-related contaminant?

d. The risk categories in 8.2 Risk Management Considerations appear to understate the real ecological risks. Are these risk categories considered to be linear?

e. How are the follow-up actions in Table 8-2 related to the risk categories?

f. What is the purpose of section 8.3.3, Seafood Contamination Levels? This is an important issue that requires more comprehensive and in-depth discussion. It belongs in a human health risk assessment, not in an ecological risk assessment.

3. Recommendations:

A more comprehensive discussion of the conclusions is needed. The discussion must tie all the related findings together. It is not surprising that the highest risks are near Seavey Island. Is there sufficient support to conclude that there are no risks at other locations? If so, it should be discussed in this section. Even though lead is a highly toxic chemical in mammals, other toxic chemicals are present which should be considered.

The levels of risk identified (negligible, low, intermediate and high) determine the risk management decisions. Risks identified as high, intermediate and low are considered in the clean-up recommendations. It is important to show clearly how the ranking decision is assessed. The analysis in the review of the risk characterization section will illustrate other specific difficulties with the conclusions.