N00129.AR.001138
NSB NEW LONDON
5090.3a

UNITED STATES ENVIRONMENTAL PROTECTION AGENCY
REGION 1
1 CONGRESS STREET, SUITE 1100
BOSTON, MASSACHUSETTS 02114-2023

July 19, 2005

Mark Evans, Remedial Project Manager
U.S. Department of the Navy
Naval Facilities Engineering Command
Northern Division
10 Industrial Highway
Code 1823, Mail Stop 82
Lester, PA 19113-2090

Re: Thames River toxicity resampling

Dear Mr. Evans:

EPA extensively reviewed the data from the Thames River toxicity sampling in consultation with the developers of the toxicity test method. This letter addresses the specific question of how to use the existing data, and whether there is a need for further testing. In addition, EPA raises new concerns about how any additional test data will be evaluated with respect to laboratory controls and reference areas.

As you know, the test was impaired by the failure to meet laboratory control performance standards, which raises fundamental questions about the initial health and condition of the test organisms. Also of concern is the extreme degree of variability within replicates of control, reference, and site samples. This variability may reflect a problem with initial test organism condition, or could indicate a problem with the test protocols and procedures related to randomization, homogenization of samples, or test conditions between replicates. The Navy did not state whether a concurrent reference toxicant test was performed, so EPA therefore assumes that there was no such concurrent test. The reference test could shed light on the condition of the test organisms.

EPA is aware that the Navy has studied this test to identify the source of this problem, and that no apparent reason for these problems has been identified. Such problems are not unprecedented with the *Leptocheirus* test, but it has been the experience of EPA that a repeat test can resolve these problems. It is possible that for some unknown reason the test organisms were not in suitable condition for testing. I therefore recommend that an additional test be run using all but a few of the same sample locations. In an attempt to use as much of the existing data as possible, EPA has studied this data set and proposes several samples that can reasonably be considered to be non-toxic on the basis of this data set, and on the basis of comparison to reference locations from other recent tests. One sample is proposed for consideration as a toxic sample on the basis of this test.

The criteria for omitting certain stations from further testing as non-toxic are as follows:

- High survival, and low variability in survival. The quantitative criterion being that no more than one replicate out of ten had less than 50% survival, and most replicates had 75% to 100% survival.
- Relatively high mean amphipod weight. The quantitative criterion being that most replicates have mean weights above 1.5 mg.
- High reproduction numbers and low variability in reproduction. The quantitative criterion being most replicates having 20 or more neonates.

On the basis of these criteria, the following stations are suggested to be considered non-toxic, and could be omitted from further testing: Zone 4, Station Z4-S2 and Pier 1, Station P5.

Similarly, the data were examined to determine whether any sediments could reasonably be considered to be toxic on the basis of this test. This is a more difficult determination because many samples had an almost even split between replicates with very low survival, and replicates that would meet the non-toxic criteria. Since the reason for the variability is not known, it is unclear what an "average" condition really should be. For example, if the variability was caused by incomplete homogenization, there is a possibility that the average condition would be a "low magnitude unacceptable risk," because more homogenization would mix the poor-performing sediment with the high performing sediment. On this basis, only one sample, Station Zone 4, Z4-S6, had very consistent low survival, low mean organism weight (all below 1 mg), and low reproduction (except for a single replicate). I suggest that this sample be considered toxic and omitted from further testing.

It should be understood that if these determinations are accepted, the designations as toxic or non-toxic would be taken, without statistical analysis, into the findings of the ecological risk assessment. It is not useful to perform additional statistical analysis of this data set owing to the extreme variability. This leads the discussion to the additional concern regarding interpretation of the re-test data.

The Sampling and Analysis Plan (SAP) contains a disconnect in logic that was not noted in previous reviews. The SAP states that a determination of "low magnitude unacceptable risk" or "high magnitude unacceptable risk" would only be made for samples with significantly lower survival than both the laboratory control and the Thames River Reference Area survival, *and* survival less than or equal to 70% of mean Thames River Reference Area survival.

There are a number of problems associated with this approach. The most egregious problem concerns the use of the "70% of Reference" approach. It is very possible that the Thames River Reference Area could, in a re-test, have a mean survival of 70%. If this were the case, a site station would only be considered "unacceptable" if it had survival of 50% or less. This is

tantamount to using an LC50 as a reference toxicity value (RTV) and is completely inconsistent with EPA's approach to ecological risk assessment.

The 70% criterion appears to be an over-extension of the statement in the SAP that in large-scale tests, over various seasons and over large areas, *Leptocheirus* has been found to have a survival of 68% or more. This statement from the SAP has not be traced to its source as part of this review, but it appears to be a statement supporting the broad applicability of the test over many types of conditions, not an endorsement of 68% reference control as a desirable outcome. Regardless of the source and purpose of this statement, it does not provide any justification for setting an acceptable risk level at 50% survival.

I suggest that the laboratory control be used as little as possible in test interpretation. The laboratory control is used to evaluate the test organism condition, test water, food, temperatures, techniques, and other variables that can affect the test outcome and confound any findings pertaining to chemical contamination. The data analysis should instead focus on the Thames River Reference Area only. To account for variability and overall degradation within the lower Thames River, no Thames River Reference Area samples are proposed for omission from the re-test. As a side note in support of re-running the test, the authors of the manual for this test indicated that 80% survival in the laboratory control was, in itself, a compromise position. The developers of the test felt that 90% would be a reasonable criterion for test acceptability based on their extensive testing, but they made a concession to use 80% in order to minimize the need to have repeat tests in broad application over diverse areas.

The approach presented in the SAP, on closer inspection, represents significant deviations from accepted and established protocols presented in the EPA manual for this test. It also completely fails to adequately address the potential for growth or reproduction effects. The following simplified approach is suggested for evaluating re-test data:

- Establish that the test conditions and organism health was acceptable by evaluating the laboratory control with respect to test acceptability criteria.
- Provided the test is acceptable, follow the statistical methods provided in the EPA toxicity test manual to determine whether any site station means show a statistically significant difference in survival, reproduction, or growth from the Thames River Reference Area stations. A more detailed discussion of statistical approaches for this type of test is found in *Methods for Measuring the Toxicity and Bioaccumulation of Sediment-associated Contaminants with Freshwater Invertebrates, Second Edition*, U.S. EPA, Office of Water, March 2000 (EPA/600/R-99/064).
- Any station that shows a significant difference in survival, growth, or reproduction will be considered to pose some level of risk, and these findings will be carried into the overall ecological risk assessment for the site.

I look forward to working with you and the Connecticut Department of Environmental Protection to protect the environs of the Naval Submarine Base. Please do not hesitate to contact me at (617) 918-1385 should you have any questions or wish to arrange a meeting.

Sincerely,

Kymberlee Keckler, Remedial Project Manager
Federal Facilities Superfund Section

cc: Mark Lewis, CTDEP, Hartford, CT
Melissa Cokas, NSBNL, Groton, CT
Bart Hoskins, USEPA, Boston, MA
Jennifer Stump, Gannett Fleming, Harrisburg, PA