Response to 15 February 2008 Letter from
Congressmen Mike Thompson and Dennis Rehberg

To help organize our reply, we restate your numbered comments and respond to each in succession.

1. The study acknowledges that “up to five Army light” tugboats participated in “several” Project SHAD tests, but it claims that complete personnel rosters were never found by the Department of Defense (DOD) or by IOM. According to SHAD veterans, the rosters were provided. For instance, a roster of personnel involved in the 1965 Shady Grove test, approximately 106 participants, was provided to IOM and confirmed by DOD. With the rosters identified and made available, we would expect the personnel to be considered in the study.

Response:

All individuals identified on rosters of tugboats, as verified by the DoD, were included as eligible participants in our study. Our records show that DoD provided us with a list of 155 names of personnel who served on the tugs. The list was found to include duplicated names. When these duplicates were removed, the list included 104 unique individuals. Thirty-four of these eligible participants were known to be deceased, leaving 70 individuals who were sought and invited to volunteer as participants. Of these, 49 volunteered to participate in the study and 2 others were determined later to be deceased, resulting in 36 total deaths from among this special participant list. Thus, the participation rate among survivors was 49/68, or 72%.

The personnel rosters of individuals serving on the five Army light tugs are known to be incomplete. During the period of the study the IOM staff spent many weeks working with the archivists at the National Archives, but they were unable to find additional names of persons serving on these Army light tugs during Project SHAD.

More generally, the study staff sought to identify additional SHAD veterans by attending several meetings of Veterans Service Organizations (VSOs). At these meetings, the IOM staff provided information about the study and encouraged maximum participation. In addition, advertisements about the SHAD study were placed in the American Legion’s newsletter.

The issue of inclusiveness was of paramount interest to our study staff. To enhance participation, the study staff sent two letters to the known SHAD veterans and comparison veterans encouraging them to participate in the study
survey. Key leadership in both the DoD and VA signed one of these letters jointly, and leaders of several large and relevant VSOs signed the other letter. Copies of those two letters are provided in the original report. In situations where SHAD veterans did not respond to multiple mailed requests to complete the health survey, follow-up telephone calls were made by a specialized research firm working with the IOM staff to ask for their participation by completing the survey over the phone. Notwithstanding these efforts, incomplete follow-up is always of concern in survey research, and as stated in our original report, “With an overall response rate of 53 percent, we can not be confident that we have a complete picture of the health of the Project SHAD participants or their controls” (pp. 68-69).

A further concern is the possibility of “recall bias”, which means that persons who know that they have been exposed recall and report symptoms that unexposed persons would forget or shrug off. Overall, recall bias is likely to result in data that appear to show some effect of exposure even if there is none. It is not possible to tell whether the way SHAD participants reported exposures or outcomes differed in any important ways from those that might have come from either non-respondents or participants who had died.

2. Personnel who were not exposed during Project SHAD were included in the study: a) the USS Granville S Hall (YAG 40), the Desert Test Center command and Laboratory ship, was not exposed during Project SHAD; b) the USS George Eastman (YAG 39), participated only in some Project SHAD tests and not in others. We believe the inclusion of personnel from these two ships compromises the study results. We request that IOM examine how the inclusion of sterile personnel may be affected the results.

Response:

This concern centers on the belief that participants who served on the USS Granville Hall or the USS George Eastman should have been excluded from the analysis or included with the control (unexposed) group. As explained below, additional analyses that excluded the USS Granville Hall and USS George Eastman personnel yielded results that are essentially the same as the published results of analyses that included these personnel.

As noted, the personnel who served on the USS Granville Hall and the USS George Eastman were classified in the same way as other Project SHAD participants. Participants were assigned to one of four groups, depending on their test participation and the agents used in those tests. Group A included those involved only in tests that used the simulants *Bacillus globigii* and methyl acetoacetate; group B, those who participated in tests that involved only the simulant trioctyl phosphate; group C, those who participated in any test in which active agents were used (e.g., Sarin, *Pasteurella tularensis*, etc.); and group D, those who participated in tests involving any other simulant agents. Group B did
not involve the USS Granville Hall or USS George Eastman, so will not be discussed further. Controls were selected from ships of the same type and class that did not participate in Project SHAD (see pp.19-23 in original report).

Thus, the personnel who served on the USS Granville Hall or USS George Eastman could have been assigned to different groups (A, C, or D), depending on the actual tests in which they participated. We note that a DoD fact sheet on Project SHAD states that the USS George Eastman was involved in tests of sarin spraying (Operation Flower Drum).

Our study reported on health effects associated with Project SHAD participation; we could not obtain sufficient information to assess levels of exposure to specific agents. However, in response to concerns about the inclusion of personnel who served on the USS George Eastman (GE) and the USS Granville Hall (GH) in the SHAD participation group, we have performed additional analyses with and without the GE and GH groups, and for each of these two groups separately. These analyses are affected by sample size considerations. In all, we could identify 352 GE personnel and 688 GH personnel (total 1,040). For simplicity, we removed the 3 participants who had been assigned to both GE and GH from our analyses.

The 1,040 GH and GE participants were distributed across three of the four analysis groups, as explained above. This distribution across analysis groups was far from uniform. Table 1 shows the numbers in groups A, C, and D as follows: all participants, participants minus the GE and GH groups, the control group, and the GH and GE groups. We used the same controls—that is, the entire control population available—in our analyses. As can be seen in Table 1, the GH and GE participants comprised only a very small proportion of group A, and a moderate number of group D participants. However, nearly all group C participants were from GH and GE.

Table 2A contains the results of an analysis of all-cause mortality. Results are given as hazard ratios with 95% confidence limits. The analyses with and without the GH and GE personnel yield essentially the same results, which is that we observe no difference in mortality between participants and controls. Thus there is no evidence here that the inclusion of GH and GE personnel, who may have been unexposed, “washed out” any significant differences in mortality between participants and controls.

Indeed, if anything, the analyses excluding these participants show even smaller difference between them and controls. Conversely, this suggests that GH and GE mortality might be higher than that of other participants, which does turn out to be true for the GE participants. Specifically, mortality is statistically significantly higher for GE personnel in both group A and group D, reflecting relatively large differences in participant versus control crude mortality rates: 32.6% verses 25.2%, respectively, in group A; and 29.9% verses 22.0%,
respectively, in group D. These last findings were rather unexpected, especially given the relatively small sample sizes. Because this finding is associated with post-hoc testing and is in the opposite direction of the implied hypothesis, a good deal of caution is called for in its interpretation.

Table 2B contains the results of a second type of statistical analysis of all-cause mortality. For these analyses we used the more rigorous log-rank tests, rather than proportional hazards analyses that were used in the published report and appear in Table 2A. For Table 2B, we used a Kaplan Meier survival analysis and the log rank test to minimize the number of assumptions required. These analyses were stratified for age at participation, and the results are given as chi-square values and associated p-values; a p-value less than 0.05 denotes a statistically significant difference in mortality between participants and controls. In general, these results parallel those from the proportional hazards analysis, which is reassuring. Again, the analyses with and without the GH and GE personnel yield essentially the same results, and mortality is higher for GE personnel in group D, but not in group A. Because the finding for the GE personnel in group D was unexpected, especially given the relatively small sample sizes, we did an additional analysis, adding another stratum for pay grade (an important factor in other analyses), with essentially no change in the results. Again, because this finding is associated with post-hoc testing and is in the opposite direction of the implied hypothesis, a good deal of caution is called for in its interpretation.

Table 3 contains the results of the analyses of self-reported morbidity. The measure is the physical component score (PCS) of the SF-36, a widely used measure of health and one of the primary outcome measures of the study; a higher PCS score indicates a response of better physical health than a lower score. PCS score differences from 2 to 4.9 are considered “small” in magnitude (see published report p. 45 last sentence). The entries in the table are mean values, adjusted for age, race, pay code, and branch of service. Each cell shows the adjusted values for participants (top figure) and controls (bottom), with statistically significant differences in bold. Except for group C, the first two columns are nearly identical, which indicates that there is virtually no difference in results when the GH and GE participants are excluded. For group C, the exclusion of the relatively large number of GH and GE personnel results in higher PCS scores (that is, better self-reported health) for participants. In group C, both GH and GE personnel have worse self-reported health than the remaining participants, and thus their inclusion in the original analyses does not appear to have diluted the overall comparison. It should be noted that all the differences in Table 3, statistically significant or not, are small in absolute terms. Furthermore, the PCS scores shown in Table 3 are generally in the range of the national norms for males in these age groups (see Table 10-2 in the original report, p. 46).
Table 1 Numbers of Participants and Controls for Additional Analyses Involving USS Granville Hall (GH) and USS George Eastman (GE) Personnel

<table>
<thead>
<tr>
<th>Group</th>
<th>All participants in original analysis</th>
<th>Participants excluding GH and GE personnel</th>
<th>GH personnel only</th>
<th>GE personnel only</th>
<th>All controls</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group A</td>
<td>3392</td>
<td>3296</td>
<td>53</td>
<td>43</td>
<td>3615</td>
</tr>
<tr>
<td>Group C</td>
<td>746</td>
<td>109</td>
<td>425</td>
<td>212</td>
<td>1093</td>
</tr>
<tr>
<td>Group D</td>
<td>871</td>
<td>564</td>
<td>210</td>
<td>97</td>
<td>1212</td>
</tr>
<tr>
<td>TOTAL</td>
<td>5865</td>
<td>4825</td>
<td>688</td>
<td>352</td>
<td>6790</td>
</tr>
</tbody>
</table>

Table 2A Hazard Rates (with 95% confidence intervals) Comparing Participant All-Cause Mortality with Control All-Cause Mortality for Additional Analyses Involving USS Granville Hall (GH) and USS George Eastman (GE) Participants

[Statistically significant differences (p < 0.05) are in **bold**]

<table>
<thead>
<tr>
<th>Group</th>
<th>All participants</th>
<th>Exclude GH and GE</th>
<th>GH only</th>
<th>GE only</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group A</td>
<td>1.01 (0.92 – 1.11)</td>
<td>1.00 (0.91 – 1.10)</td>
<td>1.25 (0.76 – 2.04)</td>
<td><strong>1.72</strong> (1.01 – 2.92)</td>
</tr>
<tr>
<td>Group C</td>
<td>0.90 (0.74 – 1.09)</td>
<td>0.86 (0.57 – 1.29)</td>
<td>0.85 (0.67 – 1.08)</td>
<td>1.00 (0.74 – 1.36)</td>
</tr>
<tr>
<td>Group D</td>
<td>1.06 (0.88 – 1.28)</td>
<td>0.99 (0.80 – 1.23)</td>
<td>0.94 (0.68 – 1.30)</td>
<td><strong>1.57</strong> (1.06 – 2.32)</td>
</tr>
</tbody>
</table>

Table 2B Log-Rank Chi-Squares and P-values Comparing Participant All-Cause Mortality with Control All-Cause Mortality for Additional Analyses Involving USS Granville Hall (GH) and USS George Eastman (GE) Personnel

[Statistically significant differences (p < 0.05) are in **bold**]

<table>
<thead>
<tr>
<th>Group</th>
<th>All participants in original analysis</th>
<th>Participants excluding GH and GE personnel</th>
<th>GH personnel only</th>
<th>GE personnel only</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group A</td>
<td>0.07, p = 0.79</td>
<td>0.18, p = 0.67</td>
<td>0.42, p = 0.52</td>
<td>1.33, p = 0.25</td>
</tr>
<tr>
<td>Group C</td>
<td>0.63, p = 0.43</td>
<td>0.25, p = 0.62</td>
<td>0.69, p = 0.40</td>
<td>0.001, p = 0.98</td>
</tr>
<tr>
<td>Group D</td>
<td>0.08, p = 0.78</td>
<td>0.21, p = 0.65</td>
<td>0.04, p = 0.84</td>
<td><strong>5.27, p = 0.02</strong></td>
</tr>
</tbody>
</table>

5
Table 3 Adjusted SF-36 Physical Component Scores* for Participants/Controls for Additional Analyses Involving USS Granville Hall (GH) and USS George Eastman (GE) Personnel

[Statistically significant differences (p < 0.05) are in bold]

<table>
<thead>
<tr>
<th>Group</th>
<th>All participants in original analysis</th>
<th>Participants excluding GH and GE personnel</th>
<th>GH personnel only</th>
<th>GE personnel only</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group A</td>
<td>44.3/46.3</td>
<td>44.4/46.4</td>
<td>38.9/39.8</td>
<td>39.6/37.5</td>
</tr>
<tr>
<td>Group C</td>
<td>47.7/49.0</td>
<td>56.6/54.0</td>
<td>49.1/50.6</td>
<td>44.6/47.0</td>
</tr>
<tr>
<td>Group D</td>
<td>50.3/52.3</td>
<td>49.6/51.9</td>
<td>49.2/50.7</td>
<td>49.5/50.5</td>
</tr>
</tbody>
</table>

* A higher PCS score indicates better self-reported health.

In summary, the exclusion of Project SHAD participants assigned to the USS Granville Hall and USS George Eastman from Project SHAD participants made almost no difference in the comparisons of participants vs. controls in either all-cause mortality or PCS. We could find no evidence that the inclusion of GH and GE personnel obscured any differences in either health or mortality between SHAD personnel and controls. We found evidence that GH and GE personnel had worse self-reported health, and that GE personnel in group D (potential exposure to simulants only) had statistically significantly higher, not lower, all-cause mortality than comparable controls.

3. The health records of deceased Project SHAD Technical Staff, who may have died as a result of health effects stemming from exposure to Project SHAD, were not examined. We would like you to determine if the cause of death information for those individuals is available and measure what impact that information would have on the results of the study.

Response:

All readily available information about deaths of Project SHAD Technical Staff was included in our analysis. We tried to find records of all deaths among the technical staff and other participants and controls by using three sources of death information: the VA’s Beneficiary Identification and Records Locator System (BIRLS) death file, the Social Security Administration’s master death file, and the National Death Index (NDI), a record system developed by the National Center for Health Statistics and maintained since 1979. Cause of death came solely from the NDI. These sources together should identify all deaths since 1979 among SHAD participants and controls except in the unlikely event that a veteran lost touch with both the VA and Social Security and slipped through the national death certification process.

The National Death Index does not contain, and therefore could not provide, fact and cause of death for deaths that occurred before 1979 (“early” deaths); we therefore used other sources to determine fact of death for the early deaths. As
seen in Table 9-1 of our original report (p. 40), these early deaths (“date or fact of death only”) constituted less than 5% of the total subjects. However, they do constitute a sizable proportion of the total number of deaths. Among participants and controls 19.8% and 20.8%, respectively, of deaths were fact of death only without cause of death known. Thus, for the analysis of all-cause mortality (“fact of death”), we have more complete data than for cause-specific (“cause of death”) analyses. In addition we understand that although the causes of the early deaths were not available for the IOM analysis, the VA is conducting a separate mortality study that will include the causes of death for those who died before 1979. Of note, the percentage of personnel who experienced an early death, 3 out of 104 or 2.9%, is nearly the same for the light tug personnel as for the other SHAD participants and controls. A further mortality follow-up could be done in a few years, when the number of deaths may be large enough to show an effect not now evident.

4. The study failed to account for the job and duty assignments of various personnel on board the ships, which resulted in different levels of exposure. Consideration should be given to the fact that personnel had different levels of exposure during training and testing to multiple weapons, experimental vaccines, trace elements, simulants, and decontamination agents. These considerations should be factored in to gain the most accurate results.

Response:

Although we considered taking into account potential variation in exposure within personnel on board ships used in Project SHAD, we found very limited data available relevant to exposure. Data relating to individual exposures were not systematically collected for the agents used in the SHAD tests. During some SHAD tests, swab samples were taken inside respirator masks to evaluate their effectiveness, but to our knowledge no human toxicology studies were conducted during SHAD. We did not find any documents indicating that blood, urine, or other biological specimens were collected from the participants to determine individual internal dose levels relating to any of these test agents.

Location on the ship was considered as a possible proxy for estimating actual exposure levels. Environmental monitors were, in fact, used to determine the level of agent penetration throughout the ship, but the data they produced are still classified because they could reveal whether and where vulnerabilities may exist in US naval vessels. Theoretically, occupational codes (e.g., boilerman) could be related to where a person was located on a ship. However, the IOM staff researchers were not confident that occupational codes would be a valid basis for estimating exposure even if the agent levels in various ship compartments were declassified. They reached this conclusion because (1) it was not possible to know the actual location of participants aboard the ships during the testing; (2) sailors were likely to be moving throughout the ship during the testing; and (3) some
sailors might have been performing duties that were different from those suggested by their occupation code. Exposure to decontaminants and receipt of experimental vaccines would be even less well documented, we believe, and so even more difficult to ascertain and study.

Our analyses thus took into account factors that could be expected to affect the mortality and morbidity of SHAD participants and for which data were available. In most of the mortality analyses, we adjusted for age, race, service branch, and pay grade. In many of the morbidity analyses, we accounted for these factors as well as smoking, drinking, and body mass index. In short, the basis for determining study participants’ exposure classification was central to the study design. Our decision was to use “test participation” as the criterion for establishing “exposure.” The lack of explicit information about individual exposures does introduce the possibility of misclassification, such that individuals in the participation group may not have been actually exposed to any agents, but we see no feasible way to get better exposure data. However, the lack of any substantial overall effect, statistically significant or not, argues against any “hidden” effect of much size.

5. The description of the tests performed does not reflect the way in which the SHAD test was actually conducted. SHAD veterans must be consulted to ensure that any existing misconceptions in the IOM study are rectified.

Response:

We did our best to incorporate input from SHAD veterans, including public testimony at committee meetings. If there is concern that we have not accurately described some features of how the SHAD tests were actually conducted, we would welcome the opportunity to meet to clarify these points. We also would welcome the opportunity to discuss your concerns regarding our review of classified material connected with Project SHAD.

To summarize, the IOM study did not reveal major differences in health outcomes between the SHAD participants and the comparison group of service personnel. We hope that those veterans worried that their participation in Project SHAD may have affected their health will find some reassurance in these results. We emphasize that this study, like any other observational epidemiological study, cannot prove that health conditions of SHAD participants are unrelated to participation in Project SHAD. Although additional studies may be helpful, without data on individual exposures it will be difficult to demonstrate a linkage between health outcomes among the SHAD veterans to these SHAD tests.

There are several intrinsic limitations to this kind of epidemiological study—both in the initial study as well as in the reanalyses conducted. These include the relatively low response rate, the possibility for recall bias, the general shortcomings of health
survey research in determining morbidity, the lack of a priori hypotheses, low sample sizes for some important outcomes, and opportunities for finding statistical significance by chance alone. Despite these limitations, we believe that the IOM’s epidemiological study was executed in a scientifically valid manner.

Frederick J. Erdtmann  
Director, Medical Follow-up Agency  
Institute of Medicine  
August 11, 2008