Analysis demonstrating that no designation of an "adequate confidence level" can support the Ballast Water Report's conclusion that five treatment types met the IMO D-2 standard but none met the 10x IMO D-2 standard

Given the test results in the reports that the Ballast Water Panel determined were reliable, there is no minimum confidence level that one could select that would support the conclusion in the Report, *i.e.* that five types of treatment systems met IMO D-2 but none met 10x IMO D-2. That's because the spread in ability demonstrated by the test results for these five treatment systems—given the volumes analyzed and the organism counts reported—is such that if the system with the weakest results is found to meet IMO D-2, then one or more of the systems with the strongest results must be found to meet at least 10x IMO D-2. Two lines of analysis demonstrate this.

Analysis 1. Calculating the highest confidence level that will support a finding that five types of treatment systems met IMO D-2, and using this confidence level to determine whether any treatment systems met 10x IMO D-2.

First, for the treatment system with the weakest results, we calculated the statistical confidence that the mean density of the entire discharge was below the IMO D-2 limit, based on the reported organism count and the volume analyzed for the trial with the weakest results (assuming, as usual, that the organisms followed a Poisson distribution, that the sampling was adequately representative and that the analytical methods were accurate).[1] If the Panel's analysis had been based on a statistical analysis of confidence, then the minimum confidence level used could not have been higher than this calculated confidence level (if it was, then the Panel could not have concluded that this treatment system met IMO D-2). We then applied this calculated confidence level to the test data for the other treatment systems, to determine whether any of these data sets demonstrate compliance with 10x IMO with at least that level of confidence. It turns out that several do. These calculations are described in Table 1.

[1] The method used for calculating the statistical confidence level is described in the attached Appendix.

Table 1. Calculation of the highest confidence level that could have been used in the Panel's analysis of IMO D-2 compliance; and application to the question of 10x IMO D-2 compliance

1. Calculating the highest required confidence level used that could have been used in the analysis of IMO D-2 compliance

In the test results for the NEI treatment system, the highest concentration of >50 µm organisms reported in any trial is 7/m^3, which corresponds to a count of 2 organisms in the 0.285 m^3 volume analyzed.[2] Assuming a Poisson distribution, these results provide a confidence level of 54% that the mean concentration of the discharge is less than 10/m^3 (*i.e.* meets the IMO D-2 standard).[3] Since the Panel concluded that the Deoxygenation + Cavitation treatment type, represented by a single treatment system, the NEI system, met the IMO D-2 standard, the Panel could not have required a minimum confidence level greater than 54% in determining compliance with standards.

2. Applying the calculated confidence level to other data for >50 µm organisms

In the trials of the Ecochlor treatment system, the highest concentration of >50 µm organisms reported in any trial is 0.3/m^3, which corresponds to a count of 1 organism in this size class in the 3 m^3 volume analyzed in each trial. These results provide a confidence level of 54% that the mean concentration of the discharge is less than 0.6/m^3 in every trial, and therefore meets the 10x IMO D-2 standard of less than 1/m^3. Thus, the Panel's conclusion—that NEI met the IMO D-2 standard and that Ecochlor did not meet the 10x IMO D-2 standard for this organism class—cannot be correct.

3. Applying the calculated confidence level to data for 10-50 µm organisms

The Ecochlor, Peraclean and BalPure treatment systems reported concentrations of <0.1/mL in the 10-50 µm organism class in all trials,[4] corresponding to an organism count of 0 in the 9 mL analyzed in each trial. These results provide a confidence level of 54% that the mean concentration of the discharge is less than 0.086/mL, *i.e.* meets the 100x IMO D-2 standard of >0.1/mL. Thus, the Panel's conclusion, that NEI met the IMO D-2 standard for the >50 µm organism class and that no systems met even the 10x IMO D-2 standard for the 10-50 µm organism class, cannot be correct.

[2] Three replicate 95 L samples (*NEI Report* pp. 38, 39, 42, 55, 59).
[3] See the attached Appendix for the confidence level calculation.
[4] <0.1/mL was reported in the 10 land-based trials of the Ecochlor system, and the 12 land-based trials and 5 shipboard trials of the Peraclean system. The BalPure system reported <0.1/mL in 10 of 10 land-based trials conducted at one test facility, and 0/mL in 5 of 5 land-based trials at another facility.

Analysis 2. Direct comparison of the volumes analyzed and concentrations reported.

We also directly considered the range of volumes analyzed and reported organism concentrations. The logic we applied is: if a treatment system whose weakest trial yielded a concentration of C based on analysis of a volume V was found to have met IMO D-2, then any treatment system whose weakest trial yielded a concentration less than or equal to C/10 when analyzing a volume of at least 10V should have been found to have met 10x IMO D-2. The pair of systems shown in Table 2 meet this condition: if NEI meets IMO D-2 then Ecochlor must at least meet 10x IMO D-2.

Table 2. Concentrations and volumes analyzed for two treatment systems

Treatment System	Volume Analyzed (m^3)	Concentration (m^{-3})
NEI VOS	0.285	7
Ecochlor	3	0.3

Conclusion

To be clear, we *are not saying* that we believe the Panel's analysis used a 54% confidence level to determine the validity of test results. We don't believe that any statistical analysis of confidence was conducted at all, which is what the SAB Report's methods indicate in stating that treatment systems were scored "in accordance with the approach suggested in the IMO G8 guidelines" and "the more detailed ETV Protocol" (Final Report, page 31). Both the IMO G8 and ETV protocols assess treatment systems simply in terms of whether the concentrations determined from the sample counts are less than the concentration limits in the standards,[5] without any statistical assessment of whether the resulting confidence that the concentrations in the entire discharge are less than the standards is above some required minimum level.

However, some Panel members have suggested that the Report's conclusion was based on an assessment of statistical confidence. We did this analysis to determine whether this is possible, that is, whether there is an assessment of statistical confidence that could support the Report's conclusion. This analysis shows that there isn't.

[5] *E.g*, see IMO G8 Guidelines, Annex 4, § 2.3.5 and § 4.7.

Public Comment submitted by Dr. Andrew Cohen 8/10/2016

Appendix: How we calculated statistical confidence levels

The website www.statsdirect.com/help/default.htm - rates/poisson_rate_ci.htm gives the formula for the upper confidence limit for the estimated incidence rate of an event modeled with a Poisson distribution. For ballast water tests, the "event" is the observation of a live organism in the sample, and the "incidence rate" is the organism concentration calculated from these observations, which is equal to the organism count divided by the volume of ballast water analyzed. The website gives the formula in terms of Y, the number of observed events (in the ballast water case, the organism count), with

> the upper confidence limit for the number of events equal to 1/2 times the Chi-square quantile for the confidence limit and the relevant degrees of freedom

and

> the degrees of freedom for the upper confidence limit equal to 2 x (Y+1)

We calculated Chi-square quantiles using the online calculator for the Chi-square distribution at:

> www.solvemymath.com/online_math_calculator/statistics/quantile_calculator.php

What we want to know is, for the weakest trial of the treatment systems that the Panel determined met the IMO D-2 limits, what confidence do the trial results give us that the average organism concentration in the entire discharge met the IMO D-2 limits? This is obtained by back-figuring. The weakest trial was a trial of the NEI VOS system that analyzed three 95-L replicates for organisms >50 µm in size and reported a concentration of 7 organisms/m^3 (*NEI Report*, Appendix K). To meet an upper confidence limit of <10 organisms/m^3, the organism count would have to be no more than 2.85 organisms (10 organisms/m^3 x 3 replicates x 0.095 m^3/replicate). By the statsdirect.com formula this corresponds to a Chi-square quantile of no more than twice this, or 5.7. The reported concentration of 7 organisms/ m^3 corresponds to a count of 2 organisms in the 0.285 cubic meters analyzed; the relevant degrees of freedom are thus 2 x (2+1) = 6. Using the online Chi-square quantile calculator we entered 6 degrees of freedom and then tried different p values until we homed in on a value that yielded a Chi-square quantile of just under 5.7 (this is the back-figuring part). That value is 0.54, yielding a Chi-square quantile of 5.68. Thus, the NEI VOS test result of 7 organisms/m^3 derived from the analysis of three 95-L replicates gives us 54% confidence that the average concentration in the entire discharge is <10 organisms/m^3.