

From: Andrew Cohen [REDACTED] >
Subject: Re: A response to Dr. Indy Burke's letter
Date: March 30, 2016 5:27:53 AM PDT
To: Peter S. Thorne [REDACTED]
Cc: Judy Meyer [REDACTED] >, Ernest Benfield <[REDACTED]>, Allen Burton [REDACTED] >, William Clements <[REDACTED]>, Lisa Drake <[REDACTED]>, Charles Haas [REDACTED], Thomas La Point [REDACTED] >, Edward Lemieux [REDACTED] >, David Lodge [REDACTED], Kevin Reynolds <[REDACTED]>, Jim Sanders [REDACTED] >, Mario Tamburri [REDACTED] >, Nick Welschmeyer [REDACTED] >, Amanda Rodewald [REDACTED] >, Peter Chapman [REDACTED] >, Robert Diaz [REDACTED] >, JoAnn M. Burkholder [REDACTED] >, L. Conquest [REDACTED], Indy Burke [REDACTED] >, Fred Dobbs [REDACTED] >, Wayne Landis <[REDACTED]>, Christopher Zarba <zarba.christopher@epa.gov>, Thomas Carpenter <Carpenter.Thomas@epa.gov>, Deborah Swackhamer [REDACTED]

Dear Dr. Thorne and Ballast Water Panel members,

I'm writing in response to Judy Meyer's March 7 letter. If Judy's letter is provided to the current SAB members, please provide my response to them also.

Dr. Meyer's letter is a reply to the claim that the main conclusion of SAB's Ballast Water Report is false, a claim made by one-third of the members of the Ballast Water Advisory Panel. The Ballast Water Report concluded that:

- Five types of shipboard treatment had demonstrated the ability to meet the IMO standards (also called the "IMO D-2 standards" and, more-or-less interchangeably, the "Phase 1 standards");
- No types of shipboard treatment had demonstrated the ability to meet standards that are 10 times more stringent (the "10x IMO" standards); and
- Even with improvements these types of treatment would not be able to meet standards that are 100 times more stringent (the "100x IMO" standards), and that "wholly new" types of systems would need to be developed in order to do so.

However, these Panel members say that when the data examined by the Panel are analyzed by the methods described in the report, three types of treatment systems demonstrated the ability to meet the 10x IMO standard in both organism size classes, and at least came very close to meeting the 100x IMO standard for one organism size class.

Dr. Meyer's main argument—what she refers to as the crux of the issue—is that the test

data examined by the Panel were not adequate to support a finding that any system met the 10x IMO standards. She bases this on statements in the report indicating that *"current testing methods do not provide the resolution required to conclude that 10X standards can be met"* (her emphasis).

I'll first discuss the statements of this sort in the report, since they are the focus of Dr. Meyer's argument. I'll then discuss a few other points, including the report's conclusion about the 100x IMO standards.

(1) Where did the statement—that "current testing methods do not provide the resolution required to conclude that 10X standards can be met"—and other similar statements in the Ballast Water Report, come from?

As we wrote in the "Corrections Needed" attachment to our March 2, 2015 letter to the EPA Administrator, there are nine statements in the Ballast Water Report that assert or imply that it is impossible to determine from the available test data whether or not any of the treatment types demonstrated the ability to meet the 10x IMO standard. Dr. Meyer quotes five of these statements in her letter. These statements are not so much wrong as they are incomplete in failing to specify the fundamental assumptions underlying them, and are therefore meaningless as they stand.

For any analytical procedure used to assess concentrations of organisms in water, whether a set of test results provides an adequate degree of statistical resolution relative to a standard depends on five factors: the minimum acceptable level of confidence that the average organism concentration in the discharge is below the standard, the statistical distribution used to estimate the level of confidence, the volume of water analyzed in each trial, the organism count in each trial, and the number of trials. There is near-universal agreement that the Poisson distribution is the appropriate choice for assessing confidence levels in ballast water testing, so presumably the Poisson distribution is assumed by the nine statements in the report. Then, for any analytical procedure, and any given count, the confidence level is raised if the volume analyzed per trial or the number of trials is increased.

To determine whether a set of test results adequately demonstrates compliance with a standard, one must know the minimum acceptable level of confidence that is assumed, the volume analyzed, the number of trials, and the counts. The volumes, trials and counts can be determined from the test reports, but the minimum acceptable level of confidence must be specified. What this level should be is not a scientific or technical question but a policy one, based in part on one's view of the optimal tradeoff between false positives (passing some treatment systems whose sample concentrations are lower than the standard but whose average discharge concentration is higher than the standard) and false negatives (failing some treatment systems whose sample concentrations are higher than the standard but whose average discharge concentration is lower than the standard). The critical point is that the acceptable level of confidence (or, in the language of Dr. Meyer's main quote, "the resolution required") must be

specified for the statement to have any meaning, and none of the nine statements specify the acceptable level of confidence or requisite resolution on which they are based. Nor is this specified anywhere in the report.

So, where did these nine statements come from? Most of these statements occur within the report's statistics discussion in Chapter 3 and Appendix C, while the others appear to be paraphrases of the statements in the statistics discussion. The report's statistics discussion closely follows similar discussions in two EPA documents, Report #EPA/600/R-10/031 ("Density Matters") and Report #EPA/600/R-10/146 ("ETV Protocol"), but the statements themselves do not appear in the EPA source documents. Indeed, the statements stand out as anomalous categorical statements within a discussion that is otherwise an explanation of how to conduct statistical analyses that are relevant to ballast water sampling, with examples given in contingent form along the lines of "if you make these assumptions, here is what the analysis would look like."

Of the nine statements, only one includes references. This statement appears twice in the report, in Chapter 3 and Appendix C, and is quoted by Dr. Meyer:

"At present, confirmation of the Phase 1 standard (< 10 protist-sized organisms mL^{-1}) represents the practical limit that can currently be achieved by testing facilities in the U.S. (e.g., MERC 2009a, 2010a, 2010b; Great Ships Initiative 2010)."

There are two things to note about this statement. The first is that "practical limit" has a special meaning here. As the earlier discussion makes clear, whatever resolution/confidence level was then being achieved by U.S. test facilities, it could be increased by analyzing a few more mL of water in each trial, and/or by increasing the number of trials. Obviously, neither of these is physically impossible, though they would involve some additional cost. Thus, "practical limit" in this case apparently means something along the lines of: this is the best we can do because the equipment manufacturers funding these tests (or the agencies, if government agencies provided some funding) don't want to pay for more extensive analyses or a greater number of trials than are needed to demonstrate compliance with the IMO/Phase 1 standards.

The second thing to note is that this statement is made in regard to U.S. test facilities. The three treatment systems whose test data demonstrate that they met the 10x IMO standard were tested in a Dutch facility, which analyzed larger volumes per trial, conducted more trials, and (in the tests of these three systems) yielded lower counts than in the treatment system tests conducted by U.S. facilities. These tests thereby achieved much higher levels of confidence, relative to any given standard, than the U.S. tests.

Most of the other statements listed in the "Corrections Needed" attachment, including the statement quoted and italicized by Dr. Meyer, appear to be variations on this statement, but with the wording limiting it to U.S. facilities left out. In summary then, the

statements in the report asserting that available test methods don't have the required resolution to determine if a treatment system meets the 10x IMO standards are (1) meaningless unless their assumptions are specified, especially the assumptions about what level of resolution/confidence is required; (2) appear to be based on limits due to the amount of funding provided for the tests and not on physical limits, since there are straightforward ways to increase confidence levels, though at some additional cost; and (3) appear to be limited to conditions at U.S. test facilities (although the relevant phrase was left off of many of the statements), and thus are not applicable to the three treatment systems that demonstrated compliance with 10x IMO standards, since they were tested at a non-U.S. facility in tests with much higher confidence levels than was achieved in tests at the U.S. facilities.

(2) Other Issues

As discussed previously with the Panel members (Andrew Cohen and Fred Dobbs letter to Panel members, Feb. 2, 2015), because of the spread in capabilities between treatment systems revealed in the test data, there is no statistical analysis possible that would support the Ballast Water Report's conclusion that five types of treatment systems demonstrated the ability to meet the IMO standards but none demonstrated the ability to meet the 10x IMO standards. Dr. Loveday Conquest, who chaired the subcommittee that drafted the statistics sections of the report, verified this. This analysis can be downloaded at www.bioinvasions.org/sab ("Statistical Analysis").

In addition, as also noted in the letter to Panel members, the actual method for assessing the test data described in the Ballast Water Report (in Chapter 4) does not include or mention any statistical analysis of results and rejection of results if they don't meet some specified minimum level of confidence. Rather the assessment method described is based on whether the sample concentrations meet the standard. For example, the report states that treatment systems were scored on whether, based on the test data, they had been "demonstrated to meet this standard in accordance with the approach suggested in the IMO G8 guidelines" (Ballast Water Report, page 31 and Table 4-1). The G8 guidelines don't mention any statistical analysis or minimum level of confidence, but rather state that a system will be found to meet a standard if it is tested in a set of trials as described in the guidelines and the concentrations in the samples meet the standard in each trial.

The 10x IMO standard for the 10-50 μm organism group is $<1/\text{mL}$, and the 100x IMO standard is $<0.1/\text{mL}$. The test results for three treatment systems considered by the Panel are as follows:

Ecochlor treatment system

Land-based Trial #1	$<0.11/\text{mL}$
Land-based Trial #2	$<0.11/\text{mL}$
Land-based Trial #3	$<0.11/\text{mL}$
Land-based Trial #4	$<0.11/\text{mL}$

Land-based Trial #5	<0.11/mL
Land-based Trial #6	<0.11/mL
Land-based Trial #7	<0.11/mL
Land-based Trial #8	<0.11/mL
Land-based Trial #9	<0.11/mL
Land-based Trial #10	<0.11/mL

BalPure treatment system

Land-based Trial #1	<0.11/mL
Land-based Trial #2	<0.11/mL
Land-based Trial #3	<0.11/mL
Land-based Trial #4	<0.11/mL
Land-based Trial #5	<0.11/mL
Land-based Trial #6	<0.11/mL
Land-based Trial #7	<0.11/mL
Land-based Trial #8	<0.11/mL
Land-based Trial #9	<0.11/mL
Land-based Trial #10	<0.11/mL
Shipboard Trial #1	<0.2/mL
Shipboard Trial #2	<0.2/mL
Shipboard Trial #3	<0.2/mL
Shipboard Trial #4	<0.2/mL
Shipboard Trial #5	<0.2/mL

PeraClean treatment system

Land-based Trial #1	<0.11/mL
Land-based Trial #2	<0.11/mL
Land-based Trial #3	<0.11/mL
Land-based Trial #4	<0.11/mL
Land-based Trial #5	<0.11/mL
Land-based Trial #6	<0.11/mL
Land-based Trial #7	<0.11/mL
Land-based Trial #8	<0.11/mL
Land-based Trial #9	<0.11/mL
Land-based Trial #10	<0.11/mL
Land-based Trial #11	<0.11/mL
Land-based Trial #12	<0.11/mL
Shipboard Trial #1	<0.11/mL
Shipboard Trial #2	<0.11/mL
Shipboard Trial #3	<0.11/mL
Shipboard Trial #4	<0.11/mL
Shipboard Trial #5	<0.11/mL

The sample concentrations obviously meet the 10x IMO standard of <1/mL in every single trial. Furthermore, simple inspection of these results suggests that our confidence

that the average discharge concentration is below 1/mL should be quite high. For the Ecochlor system, for example, there were 10 trials with 9 mL of treated discharge examined in each trial; no living organisms in this size class were observed in the samples in any trial. If the average discharge didn't meet the 1/mL standard, we would have expected to observe at least 9 organisms in the samples from each trial, and at least 90 organisms in all, instead of zero organisms.

Similarly, for the BalPure system we would have expected to observe at least 9 organisms in the samples from each land-based trial, and at least 5 organisms in the samples from each shipboard trial, or at least 115 organisms in all, instead of zero. For the PeraClean system we would have expected at least 153 organisms in all, instead of the zero organisms actually observed. Thus it is highly unlikely that these systems' average discharge concentrations did not meet the 10x IMO standard for this organism class.

For each individual trial in which 9 mL were analyzed and no organisms were observed (that is, for every trial listed above except for the shipboard trials of the BalPure system, in which 5 mL were analyzed), the level of confidence (based on the Poisson distribution) that the average discharge concentration is <1/mL is 99.98%. We could also calculate a confidence level for each set of trials, by using the summed Poisson method of Miller (2011), discussed in Appendix C of the Ballast Water Report; these confidence levels would be extraordinarily high.

These data and confidence levels further show that the statements quoted by Dr. Meyer—asserting that the test methods cannot provide adequate confidence/resolution to demonstrate compliance with the 10x IMO standards—are simply wrong, unless they assume that the minimum acceptable level of confidence per trial is something more than 99.98%.

Finally, these results show that these three treatment systems, at the very least, came very close to meeting the 100x IMO standard of <0.1 organisms/mL in every trial, and may indeed have met it in every trial. Almost certainly, the average concentration over all the trials for each system met the 100x IMO standard. Thus there is no basis for the Ballast Water Report to have concluded that it was unlikely or impossible for these systems to meet the 100x IMO standard, that they couldn't meet even with reasonable or feasible modifications, and that wholly new types of treatment systems would be needed to meet the standard. Because the organism concentrations were below detection limits in every single trial, there is simply no evidence to support any conclusion stating or suggesting that these systems were incapable of meeting any particular standard.

One last point should be made about another statement in Dr. Meyer's letter, not because it bears on the ability of these treatment systems to meet standards, but

because it provides part of the explanation of how false statements in the report about the treatment systems' abilities went undetected and uncorrected by the Panel. Dr. Meyer wrote:

"The data and review procedures were presented to the entire Panel on several occasions, over several months, and there were no objections regarding these methods or conclusions."

This is incorrect. The data were never "presented" to the Panel. Indeed, in the January 2011 meeting when I asked the Panel members in the subgroup that had conducted the analysis to tell us where we could find the data in the roughly 130 technical documents on shipboard treatment systems that had been provided to us, they refused to do so. Their reason for not telling us was that they didn't want to make public the names of the treatment systems which they had concluded had met the IMO standards, for fear that this would be seen as an endorsement of these treatment systems by the Panel or the EPA. They did not tell us which documents contained the data until we received the "Final Concurrence Draft" of the report on May 24, 2011; in Appendix A of this draft, which listed all the documents that had been provided to us, they marked with shading the documents that contained the test data they had analyzed. Since we had to review the 183 page draft and indicate our final concurrence or non-concurrence by May 27, the Panel had no feasible opportunity to review the test data.

Had the data actually been presented to the Panel, there is no doubt in my mind that the report's conclusions would have been different.

Sincerely,

Andrew Cohen

Director, Center for Research on Aquatic Bioinvasions (CRAB)

5994 McBryde Ave, Richmond CA 94805-1164

email: [REDACTED]

phone: [REDACTED]

[Center for Research on Aquatic Bioinvasions](#)
[The Exotics Guide](#)