The effects of possible changes to the Canadian 2-tiered assessment framework for dredged material based on outcomes of the 2006 Contaminated Dredged Material Management Decisions Workshop (CDMMD) are evaluated. Expanding on the “data mining” approach described in a previous paper, which focused solely on chemical lines of evidence, the efficacy of Tier 1 approaches (increases to the number of chemical analytes, use of mean hazard quotients, and the use of a screening bioassay) in predicting toxicity are evaluated. Results suggest value in additional work to evaluate the following areas: 1) further expanding minimum chemical requirements, 2) using more advanced approaches for chemical interpretation, and 3) using a screening-level bioassay (e.g., Canadian solid-phase photoluminescent bacteria test) to determine whether it would complement Tier 1 chemistry as well as or better than the solvent-based Microtox™ test method evaluated in the present study. 

Keywords: Dredged material management Sediment toxicity Canadian Disposal at Sea Program Chemical action lists Biological test method
Although Apitz and Agius (2013) considered potential outcomes using chemical data, they did not consider outcomes in the context of a tiered decision framework that considers both sediment chemistry and sediment toxicity. To assist with these endeavors, the work described here further explores CDMMD workshop recommendations by combining the database on sediment chemistry with data on colocated toxicity measures to evaluate how effectively various Lower Action Level Tier 1 approaches predict sediment toxicity, and to evaluate potential regulatory outcomes for a range of tiered DM assessment protocols that include both chemical and toxicological evaluation. Specifically, the present paper evaluates the performance of bioassay results similar to Canada’s current Tier 2 battery of bioassays against current chemical screening results, and the outcomes and implications of various potential changes to the Canadian chemical and toxicological assessment protocols for DaS, including the assessment of a broader suite of metal and organic contaminants. Outcomes following the application of a range of Tier 1 Lower Action Level (AL1) decision rules, including the use of mean hazard quotients (mHQs) and the addition of a screening bioassay, are also evaluated. The potential efficacy of chemical Upper Action Levels (AL2) in predicting toxicity failures will be addressed in a subsequent paper.

Glossary of terms (adapted from IMO 2009)

**Action level (AL).** A decision rule or set of decision rules that integrates findings related to multiple lines of evidence (characteristics), after comparison to their respective benchmarks, to yield a single, “overall” decision.

**Chemical action level.** Uses only chemical benchmarks or sediment quality guidelines (SQGs).

**Lower action level (AL1).** Action levels that identify levels below which there is “negligible environmental concern” in relation to disposal decisions.

**Upper action level (AL2).** Action levels used to avoid acute and chronic effects in relation to decisions about disposal at sea.

**Action list.** Comprised of a number of characteristics to be considered for measurement in the dredged material.

**Benchmark.** A point on the range of the metric (e.g., 4 mg/kg Cu, 20% amphipod mortality) that is used to identify where environmental concern may be low or high for that characteristic. These can be referred to as the “lower benchmark” and “upper benchmark.”

**Characteristic.** Ann attribute of dredged material (e.g., Cu, Hg, silt, petroleum compounds, pathogens) or a biological response to dredged material (e.g., mortality, growth, bioaccumulation).

**Metric.** A measurement that can be made on a characteristic (e.g., concentration, percent survival).

**METHODS**

Figure 1 depicts the steps followed to collect and analyze data in the present study.

**Database development**

The database was developed as described in Apitz and Agius (2013). As reported there, bioassay results were collected for subsequent analysis but had not at that point been interpreted or validated. A compact data set was generated with all the sediment records that had the minimum chemical data set designated in Apitz and Agius (2013) and also contained a full “matched” bioassay battery with results for amphipod (Ampelisca) survival, sea urchin fertilization, and Microtox™. These particular bioassays were selected because they were the closest available matches to the current Canadian protocols. Data for bioassays in the database were assigned pass or fail values on the basis of the criteria described in Table 1.

The resultant compact data set contained a broad range of sediment physical, chemical, and biological data. Data sets were reviewed to ensure that all results for a given parameter were in the same units, and anomalous data (such as nonnumerical results or impossible values such as negative concentrations) were eliminated unless they could be corrected in correspondence with relevant database coordinators. Because the objective was to generate a plausible and realistic proxy for data that might potentially be encountered, but not to draw specific risk conclusions from specific data or sites, no further data validation or quality control were carried out. The number of records with both biological and chemical data was about half of the number that were available for the chemical analysis in Apitz and Agius (2013) but was still substantial; there were 1081 records with the complete data, with a reasonable distribution around the US coastline.

**Chemical parameters of interest**

Apitz and Agius (2013) described the methods and assumptions used to select chemical parameters of interest and the range of chemical action levels tested for each parameter. Briefly, the focus was on comparing sediment data to a set of chemical ALs that might be used as AL1 values in a decision framework (although these were called “Lower Action Levels” in Apitz and Agius [2013]). The set of chemical parameters used included metals for which other regulators have established benchmarks, total PAH (defined as the sum of the US Environmental Protection Agency [USEPA] priority PAHs for which data were available), total PCB (defined as the sum of congeners 28, 52, 101, 118, 138, 153, and 180 [Webster et al. 2013]), and other organics for which both data and benchmarks were available. A broader list of other chemical analytes was included in the database for potential future use.
Lower action level (AL1) values used

To test how various changes to the DaS chemical assessment protocols would affect potential regulatory outcomes for the sediments in the database, contaminant levels were compared to hypothetical chemical AL1 for each constituent under consideration. A list of “consensus” ALs was developed by Apitz and Agius (2013) to provide a consistent basis of comparison throughout the assessment scenarios; details of their development can be found there. The consensus values were based upon the geometric mean of chemical ALs used internationally in dredging programs, or, if not available, in other sediment assessment frameworks. The ALs tested in the present study, as well as those currently used in the Canadian DaS program (with the total PCB [tPCB] value adjusted for congener-based values as described in Apitz and Agius [2013]) are described in Apitz and Agius (2013); they are also briefly discussed in Supplemental Data S1.

Different countries develop ALs on the basis of different sediment size fractions and analytical methods. Because most sediment contaminants tend to associate with the fine-grained sediment fraction, these differences could result in different pass/fail interpretations in various countries. However, overall sediment pass/fail outcomes using different AL sets with the same narrative intent (e.g., AL1, AL2) do not
differ nearly as much as outcomes using different chemical action lists and decision rules (Wenning et al. 2005; Apitz et al. 2007; Apitz 2008, 2011; Apitz and Agius 2013). The consensus AL1 values used here provide a consistent set of hypothetical ALs for the full suite of contaminants in the present study; they should not be regarded as regulatory proposals.

Use of mean hazard quotients (mHQs)

Part of the present study examines the current "one out/all out" chemical decision rule used in the Canadian DaS program, wherein the exceedance of any single benchmark, by any margin, triggers a "failure" of that sample for the entire tier. By failing a sediment for even a marginal exceedance of the AL by a single constituent, it is possible that one out/all out rules are overconservative; at the same time, by treating a sediment with a very large degree of exceedance of a contaminant in the same manner as one with only a minor exceedance, they may also be underconservative. Also, a chemical-by-chemical assessment may fail to account for the cumulative or synergistic effects of low to moderate levels of multiple contaminants, compared to slightly higher levels of a single contaminant.

The use of mHQs could potentially address these issues with one out/all out rules. Mean hazard quotients have been developed in an attempt to account for the potential effects of multiple contaminants (assuming simple additivity), as well as the magnitude of exceedance above relevant ALs. In this method, the sum of HQs for all chemicals considered is divided by the number of contaminants measured in a sample. In general, sediments with low mHQs for threshold-effects ALs (i.e., Effects-Range Low [ERL] or Threshold Effects Concentration [TEC]) are probably, but not definitely, nontoxic, and those with relatively high mHQs for probable-effects ALs (i.e., Effects-Range Medium [ERM] or Probable Effects Concentration [PEC]) are often, but not always, toxic. Samples with contaminant levels between these may or may not be acutely toxic, depending on their specific chemical characteristics.

In the present paper, in which mHQ values are used to evaluate the likelihood of nontoxicity (AL1), mHQ values were calculated on the basis of chemical AL1 values. Although a range of values and approaches were considered (see Supplemental Data), in the present paper, the mHQ pass benchmark was $mHQ_{AL1} < 0.2$.

Selecting Canadian-equivalent bioassays

Canadian DaS proponents whose sediments exceed or fail Tier 1 chemical testing and who still wish to be considered for open water disposal are required to conduct 3 bioassays to proceed through Tier 2. The 3 bioassays must include an acute lethality assay (EC 1998), and 2 of the other sublethal or bioaccumulation tests in the Canadian DaS battery, which is comprised of polychaete survival and growth in whole sediment (EC 2001), luminescent bacterial inhibition in solid phase (EC 2002), echinoid fertilization in porewater (EC 2011), and bivalve bioaccumulation in whole sediment (USEPA 1993). To "pass" Tier 2 and be considered for open water disposal, proponents must pass the lethality test and at least one of the other tests. If 1 sublethal test fails, sediments are considered "sublethally toxic," and open water disposal can only be considered with "special handling" to mitigate risk. Failure of the acute lethality test and/or both sublethal tests is considered "acutely toxic" and not considered suitable for unconfined disposal at sea.

Within the database, a range of bioassay results in a variety of units was available. These data were reviewed to determine how they compared to those used in the Canadian DaS program, and to develop scenarios that were as similar as possible to those likely to be encountered within the Canadian DaS program. Although exact matches could not be found in the original database, the first column of Table 1 lists the bioassay tests currently being used in the Canadian DaS program for which proxies were available for inclusion in the present study database. Table 1 summarizes the resulting "matched battery" of similar bioassays that was identified, developed, and used for an initial assessment of Tier 1 and 2 performance. Although the studies that fed into the database for the present work used different species, methods, and endpoints than did the Canadian DaS ones, these species, methods, and endpoints are similar enough to those used in the Canadian DaS program to be able to represent a set of plausibly similar acute and sublethal bioassays.

The most significant difference between the bioassays used in the Canadian DaS program and those in the matched database battery is associated with the Microtox™ test. For this test, Canada uses a solid-phase sediment as a test medium, whereas the data available and included in the database were generated using a solvent extract. It is recognized that these results cannot be directly compared. However, the substitution of another test, using a different organism and/or test endpoint would have rendered the matched battery even less comparable to the one used by the Canadian DaS program, and would have introduced additional complications and limitations when determining the implications of results on potential changes to the Canadian DaS assessment framework. Therefore, as a first step in evaluating the existing Canadian DaS bioassay battery within an overall assessment framework, it was decided that the solvent-based Microtox™ test using a bioluminescence endpoint would make a suitable proxy for the Canadian DaS solid-phase sediment test with the same endpoint.

Developing Canadian DaS-equivalent pass/fail criteria

For Ampelisca survival data, if survival rates were above 64.7% of control values, sediments were deemed to have passed; otherwise they failed. These pass rates were based upon those specified in the Environment Canada (EC 1998) method for situations when grain size corrections cannot be made (data for grain size and potentially confounding factors were not consistently available in the database), and with species-specific corrections based upon the pass rate for Eohaustorius washingtonianus. Sea urchin fertilization rates...
were deemed to have passed if they exceeded 70% in 100% porewater or were 70% of control values, depending upon the data available, based upon EC (2011) decision rules. Determining Microtox™ pass criteria was not as straightforward for a number of reasons. Firstly, the data downloaded from the NOAA website had no metadata on Microtox™ analyses, methods, or units. Also, detailed examination of the data revealed disparities of several orders of magnitude in the ranges of data reported from different studies and regions that fed into the database; these disparities could not be explained by contaminant levels, nor did they correspond with results for other bioassays.

Because these data are up to decades old, and many of the people involved in its generation are retired or in other jobs, extensive correspondence with the database administrator was needed to resolve Microtox™-related issues. As a result of these efforts, it was confirmed that the data in the database were based on assays conducted with solvent extracts. Also, it was determined that the data from 4 regions had been reported in different units than those in other regions. Corrected data were received and the database was updated (as were, we believe, the data in the online data source itself). The resulting Microtox™ data set was more consistent; regional data ranges were more comparable and corresponded better with contaminant level ranges.

Because the EC (2002) photoluminescent bacteria method is a solid-phase test, whereas the database results were from a Microtox™ solvent-extract test, Canada’s pass/fail criterion could not be readily applied. As a result, the present paper used pass/fail criteria for the database Microtox™ measures that were based upon values developed by Long et al. (1999).

Long et al. (1999) generated 2 new Microtox™ critical values, both based upon statistical analyses of data from a subset of the original data from which the NOAA surveys (n = 1013) that the present study draws upon were compiled. The first value (0.06 mg/mL) represents the 90% lower prediction limit (LPL) of the entire data set. The probability that a future observation from this data distribution that is less than the LPL would be more toxic (i.e., an EC50 < 0.06 mg/mL) would be 90%. Therefore, a sample with an EC50 less than 0.06 mg/mL would be extremely toxic in this test. This value is used as a Tier 2 pass/fail criterion for the Microtox™ sublethal test (Sublethal 2 in Table 1).

The second Long et al. (1999) value (0.51 mg/mL) represents the 80% LPL with the lowest (most toxic) 10% of the data values removed from the database to eliminate their influence on the distribution of the data. Samples with EC50 values < 0.51 mg/mL or > 0.06 mg/mL would be considered as moderately toxic in this test. This value is used in the present study as a conservative Tier 1 screen, as noted in the

| Database test and endpoint                          | Classification or role in this study | Organism in database | Fails if … | Canadian DaS test equivalent | Canadian DaS “fails if” criteria … |
|-----------------------------------------------------|-------------------------------------|----------------------|------------|-----------------------------|-----------------------------------|
| Amphipod survival in whole sediment                 | Acute                               | Ampelisca abdita     | % control ≤ 64.7% | Amphipod survival in whole sediment | Decrease in survival of 20% or more between test and reference sediment or 30% between test and control sediment |
| ASTM 1992                                           |                                     |                      |            | EC 1998, EPS1/RM/35          |                                    |
| Tier II Microtox™ (solvent extraction); ECS0 5-min mean | Sublethal 2                         | Allivibrio fischeri  | mg equiv sed wt ≥ 0.06 | Luminescent bacterial inhibition in solid phase | Five-minute IC50 is less than 1000 mg/kg; special considerations for grain size effects |
| Long et al. 1999                                    |                                     |                      |            | EC 2002, EPS1/RM/42          |                                    |
| Echinoid fertilization in porewater                 | Sublethal 1                          | Arbacia punctulata   | Fertilization rate in 100% PW < 70% or | Echinoid fertilization in pore water | Decrease in fertilization of 25% or more between tests sediment and control water |
| Carr 1998                                           |                                     |                      | % control 100% porewater < 70% | EC 2011, EPS1/RM/27 |                                    |
| Tier 1 screening bioassay                           | Screening bioassay                   | A. fischeri          | mg equiv sed wt ≥ 0.51 (screening criterion) | N/A                            |                                    |
| Long et al. 1999                                    |                                     |                      |            | (no screening-level bioassessment in Canadian DaS battery) |                                    |

ASTM = American Society for Testing and Materials; DaS = Disposal at Sea program; EC = Environment Canada; IC50 = % PW = pore water; SOP = standard operating procedure.

*The 4 left columns describe the bioassays and “fails if” criteria used in this study to match those used in the Canadian DaS framework. Endpoints, organisms, and roles in the study are also described. The last 2 columns describe the equivalent tests used in the Canadian DaS program.

Table 1. Canadian DaS bioassays and the “matched” bioassays used from the database*
last row of Table 1. In the present paper, Microtox™ values are compared to this more conservative value to provide a hypothetical Tier 1 "screening" bioassay in a number of scenarios tested.

It should be pointed out that both the extremely and the moderately toxic narrative intents are less conservative than those typically used for AL2 and screening, respectively, in DaS programs. However, in the absence of a comparable DaS benchmark for these endpoints, the decision was made to use published benchmarks used for sediment risk assessment to enable an initial evaluation of scenarios, including a Microtox™ screening bioassay in Tier 1.

Data analysis

A set of regulatory scenarios was developed to evaluate the efficacy (in terms of the ability to correctly predict acute and sublethal toxicity) and implications of various potential changes to the Canadian chemical assessment protocols for DaS. The scenarios tested and evaluated in the present paper include the assessment of a broader metal and organic chemical action list, the application of a range of Tier 1 decision rules, and the addition of a screening bioassay; these scenarios were then defined as test protocols. As a baseline, the records in the data set were evaluated using the current Canadian DaS protocol (i.e., the status quo), the expanded DaS chemical action list, alternative chemical rules, and the matched bioassay battery. Subsequent scenarios then changed Tier 1 protocols to examine changes in outcomes. Detailed descriptions of scenarios evaluated in the present study can be found in Supplemental Data Tables S1-2a to S1-2e. Brief descriptions of scenarios selected for discussion in the present paper can be found in Figures 2 and 3.

Table 2 summarizes the Canadian DaS tiered decision rules, considering both chemical and bioassay outcomes, and the resulting...
regulatory outcomes: unconfined open water ocean disposal, open water ocean disposal with special handling restrictions, and no ocean disposal without confinement or treatment. Briefly, sediments that pass the chemical tier (i.e., <AL1) are considered acceptable for ocean disposal without further bioassessment. If they fail the chemical tier (i.e., >AL1), to pass Tier 2 and be considered nontoxic and acceptable for unrestricted open water disposal, proponents must pass the lethality test and both of the other (sublethal) tests. If a single sublethal test is failed, sediment is considered sublethally toxic, and special handling is required for unconfined disposal at sea. Sediments that fail either the acute toxicity assay or both sublethal assays are considered acutely toxic, and open water disposal is not permitted.

**Tier 1 scenario options tested**

**Expanded chemistry.** Various Tier 1 scenarios were evaluated by comparing assessments made using the following:

1. only the regulated Canadian DaS chemical action list (Cd, Hg, total PAH, and total PCB);
The mHQs are calculated with the following formula:

$$\text{HQ}_{ij} = \left[ \frac{[C]_{ij}}{A_{Li}} \right],$$

(1)

where $[C]_{ij}$ is the concentration of contaminant $i$ in sample $j$, and $A_{Li}$ is the AL1 of interest.

For one out/all out Tier 1 scenarios, the sample fails if this value is $\geq 1$ for any analyte included in the chemical action list. In other scenarios, chemical rules based upon mHQs were tested. The mHQs are calculated with the following formula:

$$\text{mHQ} = \left( \sum_{i=1-n} \left( \frac{[C]_{ij}}{A_{Li}} \right) \right) / n,$$

(2)

where $n$ is the number of analytes in the chemical action list being considered for the sample and scenario.

Depending on the scenario, either AL1 or AL2 is used for this calculation (only AL1 results are discussed in the present paper; results for other mHQ approaches can be seen in Supplemental Data Tables S1-3a to S1-3d). To calculate mHQ values and conduct the work presented in here, a set of consensus AL values was used as described in Apitz and Agius (2013). Given the origins and objectives of the ALs used in the present study, the selection of an appropriate value for an mHQ decision rule is somewhat arbitrary; although a number of mHQ values were evaluated, only the mHQ$_{AL1} < 0.2$ results are shown in the present paper. Results based on mHQs calculated using AL2 values generally performed more poorly than the one out/all out rule in terms of Class II and III assignments (see Supplemental Data for additional information).

### Table 2. Dredged material regulatory outcomes following Tier 1 and 2 assessment with the current Environment Canada Disposal at Sea program

| Chemistry (Tier 1) | Bioassay battery (Tier 2) | Canadian DaS outcome (this paper) | Tier 2 toxicity classification (this paper) |
|--------------------|--------------------------|-----------------------------------|------------------------------------------|
| <AL1               | N/A* – No Tier 2 assessment required | Ocean disposal | NR$^b$ |
| >AL1              | Pass Pass Pass | Nontoxic | Nontoxic |
| >AL1              | Pass Fail Pass | Special handling | Sublethally toxic |
| >AL1              | Pass Fail Fail | No ocean disposal without treatment | Acutely toxic |
| >AL1              | Fail N/A N/A | Acutely toxic | Acutely toxic |

AL1 = lower action level; DaS = Disposal at Sea program.

* N/A means either test not required or results irrelevant due to other results driving decision.

$^b$ NR means not relevant; in this paper, toxicity assessed for samples <AL1 to determine false negative rate; toxicity classifications as for samples >AL1.

2) the DaS list as in 1) with the addition of 6 other metals (Cr, Ni, Pb, Cu, Zn, and As);
3) the DaS list with the addition of 7 other organics (total DDT (tDDT), the sums of dichlorodiphenyldichloethane (DDD), dichlorodiphenyldichlooroethylene (DDE), and DDT values when reported), total tributyltin (tTBT, the sum of tributyltin and dibutyltin), lindane, dieldrin, chlordane [the sum of alpha and gamma chlordane when reported], aldrin and hexachlorobenzene (HCB)]; or
4) the DaS list with the addition of all available metals and organics (full).

Mean hazard quotients chemical assessments. One out/all out chemical decision rules and mHQ chemical assessments were compared.

For each scenario, hazard quotients (HQs) were calculated for all samples and chemical constituents to be considered. For contaminant $i$ and sample $j$,

$$\text{HQ}_{ij} = \left[ \frac{[C]_{ij}}{A_{Li}} \right];$$

(1)

where $[C]_{ij}$ is the concentration of contaminant $i$ in sample $j$, and $A_{Li}$ is the AL1 of interest.

For one out/all out Tier 1 scenarios, the sample fails if this value is $\geq 1$ for any analyte included in the chemical action list. In other scenarios, chemical rules based upon mHQs were tested. The mHQs are calculated with the following formula:

$$\text{mHQ} = \left( \sum_{i=1-n} \left( \frac{[C]_{ij}}{A_{Li}} \right) \right) / n,$$

(2)

where $n$ is the number of analytes in the chemical action list being considered for the sample and scenario.

Depending on the scenario, either AL1 or AL2 is used for this calculation (only AL1 results are discussed in the present paper; results for other mHQ approaches can be seen in Supplemental Data Tables S1-3a to S1-3d). To calculate mHQ values and conduct the work presented in here, a set of consensus AL values was used as described in Apitz and Agius (2013). Given the origins and objectives of the ALs used in the present study, the selection of an appropriate value for an mHQ decision rule is somewhat arbitrary; although a number of mHQ values were evaluated, only the mHQ$_{AL1} < 0.2$ results are shown in the present paper. Results based on mHQs calculated using AL2 values generally performed more poorly than the one out/all out rule in terms of Class II and III assignments (see Supplemental Data for additional information).

Addition of Microtox™ screening bioassay. A Microtox™ screening bioassay with a very conservative pass threshold (described in Table 1) was added to some Tier 1 scenarios. When used, a sample was required to pass 1) the chemical rule (one out/all out or mHQ < X) and the Microtox™ bioassay screening criterion, or 2) only the conservative Tier 1 Microtox™ screening criterion (Microtox™-only scenarios). Various combinations of these parameters were evaluated.

Evaluating scenario results

To determine how each Tier I scenario performed in terms of predicting toxicity, scenario decision rules were applied to each sediment sample in the data set (1081 samples), and scenario outcomes were classified on the basis of their Tier 1 and Tier 2 outcomes, and whether these tiers agreed or disagreed. Although screening bioassessment Tier 1 scenarios are examined, for simplicity, Tier 1 outcomes will frequently be called “chemistry,” whereas Tier 2 outcomes are referred to as “biology” outcomes in this discussion. Scenario outcomes were classified as I to VI, as in Figure 1. Clearly, the most undesirable classifications are II and III, false negatives in which a chemical pass (and thus no subsequent bioassessment) results in toxic sediment being approved for unconfined open water ocean disposal (DaS). Class IV, a false positive in which nontoxic samples fail the chemical tier but subsequently pass bioassessment and are thus correctly approved for DaS, result in “extra” bioassessment but are ultimately successfully classified, as are Class V and VI sediments, which fail both chemical and bioassessment and are thus subject to regulatory controls.
Although a tiered program should seek to minimize false positives for cost effectiveness, this must be balanced against the overarching need to minimize false negatives to ensure that the tiered approach is sufficiently protective of the marine environment. Thus, in discussions of various assessment scenarios in the next section, effective scenarios are considered to be those that maximize the rates of Classes I, V, and VI (correct decisions) while reducing relative rates of Class II (sublethally toxic false negative) and III (acutely toxic false negative) sediments, with a secondary focus on minimizing Class IV (false positive) rates, which represent correct framework decisions made at additional Tier 2 assessment cost to proponents.

The approach taken in the present paper is based on the assumption that the Tier 2 bioassessment is the baseline measure of “truth”—that toxicity measured in field sediments evaluates whether the chemical benchmarks are appropriate. This assumption is necessary to provide some basis of comparison, but the toxicity measures themselves contain uncertainty. Bioassessment can be subject to a range of confounding factors and may be sensitive to unmeasured contaminants in the sediments; errors or misclassifications can occur in any measurement (Wenning et al. 2005; Apitz 2011). Scenarios using mHQ filters to evaluate the possibility that sediments failing a single sublethal assessment (not described here but results in Supplemental Data Table S1-3) suggest that some proportion of sediments classified as sublethally toxic (i.e., failing 1 sublethal test) could be the result of confounding factors (the authors were able to confirm that hydrogen sulfide toxicity, which has a strong confounding effect in Microtox™ bioassays done using solvent extracts, was unlikely to be the culprit because standard procedures used in the National Status and Trends Program require the removal of hydrogen sulfide prior to solvent extraction [Long et al. 1999]). The Canadian DaS tiered framework, and many others, allow for an evaluation of potential confounding factors when interpreting Tier 1 results, but data were not available for such an assessment in the present study. Thus, although substantial changes in efficacy between scenarios may be indicative of better approaches, minor differences that could be due to unique conditions in a handful of samples should be interpreted with caution.

**RESULTS**

Figures 2 and 3 illustrate the number and percentage of samples that fall in each sediment class for selected scenarios. Combined outcome classifications are based upon the agreement or disagreement of Tier 1 and Tier 2 classifications, as in Figure 1. Regulatory outcomes at the far right are shown as the percent of samples that would result in each regulatory decision after a full Tier 1 and Tier 2 assessment as per Table 2. “% Ocean Disposal” outcomes are the sum of Class I to IV outcomes. Percent “Special Handling” outcomes are comprised only of Class V samples. “% No Ocean Disposal” outcomes are comprised of Class VI samples. The results using the existing Canadian DaS decision-making rules and ALs (“Status Quo”) are shown in the first row of Figures 2 and 3. All other assessment protocols use consensus ALs, or mHQs calculated using the consensus ALs.

**Expanded chemistry**

The classification rates using “Expanded Chemistry” assessment protocols are shown just below the Status Quo (Scenario 1) in Figure 2. These scenarios evaluate different action lists, using a one out/all out rule (if a single analyte exceeds its AL, this is a Tier 1 failure) and the bioassessment protocol in Tier 2. Using the 4 Canadian DaS analytes and the consensus ALs would result in about 1% of samples being acutely toxic false negatives (acutely toxic sediments receiving a permit for unconfined open water disposal; Class III) and almost 15% sublethally toxic false negatives (Class II). About 18% of the time, Class IV false positives would have resulted, such that applicants would need to carry out Tier 2 assessment for sediments that would ultimately receive unconfined ocean disposal permits. Special handling requirements and no ocean disposal decisions (sublethally and acutely toxic true positives) would result about 5.4% and 4% of the time, respectively (Class V and VI).

The addition of metals only to the assessment protocol (Scenario 3: DaS + metals) results in a reduction of both Class II and III false negatives, but an increase in Class IV false positives. The addition of organics only to the assessment protocol (Scenario 4: DaS + organics) results in a greater reduction in Class II and III false negatives than does the addition of metals, but also a much higher rate of Class IV false positives. When both metals and organics are added to the assessment protocol (Scenario 5: Full), the Class II and III rates are lower and the Class IV rates are higher than when a single class of chemicals is added, suggesting that shorter lists of chemical constituents are less effective than the longer lists in predicting sediment toxicity, and that the addition of constituents to the Tier 1 chemical assessment substantially improves the protectiveness of the assessment protocol. The “Full” protocol does this by more than doubling the Class IV false positive rate (therefore requiring Tier 2 assessment more often). However, among these Tier 2 assessments, a large proportion of the samples will ultimately pass bioassessment and be deemed suitable for ocean disposal. Specifically, these Tier 2 assessments will pass (as Class IV) more than twice as often as they will fail and fall in Classes V or VI.

**Mean hazard quotient chemical assessments**

The bottom grouping in Figure 2 shows the results of assessment protocols using a mean hazard quotient (mHQAL1) instead of the one out/all out rule to define chemical pass or fail.

For all sets of analytes, the performance of the mHQAL1 is much better than the one out/all out rule, substantially reducing the Class II and III false negative rates, but at a cost of much higher Class IV false positive rates.
The fact that the sublethal mHQ performance for DaS + organics (Scenario 8) is worse than for DaS + metals (Scenario 7), although DaS + organics (Scenario 13) is better than for DaS + metals (Scenario 12) is better in the one out/all out rule protocols, suggests that there is still scope for work in deriving the appropriate organic ALs for a tiered assessment. The focus of the present work, however, is not to suggest the specific ALs to apply, but to demonstrate how different general approaches affect framework performance.

**Addition of a screening bioassay**

Figure 3 illustrates the classifications that would result from adding a Microtox™ screening bioassay to Tier 1. In these assessment protocols, a conservative pass/fail criterion was applied to the database Microtox™ data. This assessment was added to the 2 sets of Tier 1 chemical decision rules as follows: To pass Tier 1, the sediment would have to pass the chemical decision rule (either one out/all out or the mHQ_AL1-based rule), and pass the screening bioassay. The results allow an evaluation of whether this screening bioassay could complement Tier 1 chemistry by potentially flagging sediments that are contaminated with constituents absent from the chemical list.

The row just below the status quo on Figure 3 illustrates results when a screening Microtox™ bioassay is used alone in Tier 1 without any chemistry (Scenario 10). The rest of the grouping below the status quo in Figure 3 shows classifications following the simplest application of a screening bioassay, by adding it to a simple one out/all out approach. For all chemical action lists, the addition of the screening Microtox™ to the Tier 1 assessment increases agreement between chemical and biological screening performance in terms of predicting toxicity, but not in terms of predicting nontoxicity. Acutely toxic false negative (Class III) rates were substantially reduced; in the cases of the DaS + organics (Scenario 13) and Full (Scenario 14) list of analytes, Class III rates were reduced to 0.1%; this represents a single sample out of the 1081 examined. At the same time, the addition of the screening bioassay to the full chemical action list (Scenario 14) reduces sublethally toxic false negatives (Class II) from 6.9% to 4.4%.

The bottom Figure 3 grouping shows classifications following the combination of a screening bioassay with an mHQ in Tier 1. This assessment protocol yielded the lowest Class II (sublethally toxic false negative) rates of any scenario set, and consistently low Class III (acutely toxic false negative) rates, but the highest Class IV (false positive) rates of up to 50%.

The question of whether a screening bioassay alone would be enough, and whether chemical measures are unnecessary, is addressed by the increasingly better performance when the chemical list is expanded, even with the addition the screening bioassay. Figure 3 illustrates an assessment protocol (Scenario 10; µTox only Tier 1) in which only the screening bioassay (without chemical assessment) is carried out in Tier 1. As can be seen, the screening bioassay alone performs as well for Class III as the Status Quo (Scenario 1), which uses only the current Canadian DaS ALs, and somewhat better than the Status Quo for Class II. The screening bioassay alone (Scenario 10) performs substantially better than the DaS assessment protocol (Scenario 2 in Figure 2, which uses only consensus AL values) at minimizing Class II sediments, but performs slightly less well at minimizing Class III sediments (by 2 samples).

**DISCUSSION AND CONCLUSIONS**

**Expanding minimum chemistry**

The present study has found that for most scenarios, increasing the number of chemical analytes not only increased Tier 1 failure but also better predicted Tier 2 toxic (but not nontoxic) outcomes, suggesting that the rate of false negatives could be reduced by adding chemicals to the current Canadian DaS action list.

It is worth noting that the increased protectiveness achieved through the addition of chemical analytes comes at the cost of a (sometimes) greatly increased false positive rate compared with the current Canadian DaS framework. Among the increased number of samples subjected to bioassessment, generally well over half of these pass Tier 2 and are ultimately designated as suitable for unconfined ocean disposal. In other words, additional Tier 2 assessment does not drastically reduce the number of projects that are ultimately allowed for open water disposal. The results of this work suggest that requiring additional analyses of the chemical constituents examined in the present study is warranted because the additional information maximizes the rate at which correct “toxic” assignments are made (i.e., maximal reduction in Class II and III sediments is achieved). Also, the addition of only a few chemical constituents is needed to achieve improvements in the rate of correct decisions, and so the increased analytical burden to proponents is expected to be minimal.

The appropriate balance between maximizing the rate of correct environmental decisions and minimizing the assessment burden placed on proponents (i.e., how many false positives and negatives are acceptable) is ultimately a policy decision that environmental managers have to make when designing and applying DM assessment frameworks. When considering potential increases to the assessment burden on proponents, it may be possible to streamline assessment in other areas of the framework. For example, in the Canadian DaS context, one of the CDMMD recommendations not considered in the present paper or evaluated to date, was to streamline and reduce assessment requirements for routine, low-risk permit applications by creating a separate pre-screening assessment that is distinct from Tier 1. This separate assessment would help to focus any increased assessment efforts on permit applications associated with greater uncertainty and/or greater risk of contamination, while ensuring that low-risk projects are assessed using a streamlined process.

Of course, even with the extended chemical action list examined in the present paper, only a small fraction of the
millions of chemicals that might be found in the environment are represented (Daughton 2002); the inclusion of further organic contaminants in an action list was addressed to some extent in Apitz and Agius (2013). It has been noted that regulatory programs are skewed toward those chemicals that have the most data on human health risk, though these may not be the highest risk compounds (Wagner 1995) and that, for example, in European rivers, contaminants on priority lists could explain only a small fraction of the observed toxicity (Lahr et al. 2003; Brack et al. 2007). Generally speaking, when considering which contaminants to include in an action list, it would be worth evaluating the inclusion of so-called “emerging contaminants” in addition to the legacy contaminants typically considered by DM regulatory frameworks around the world. However, such an approach is not possible with this data-mining exercise, which relies on both extensive data on colocated sediment chemistry and toxicology and risk-relevant ALs to carry out an assessment.

Mean hazard quotient chemical assessments

Although broad differences in the mHQ approach versus the one out/all out protocol can be illustrative, they should be viewed and generalized with caution. The performance of an mHQ approach is sensitive to the chemical action list and the ALs applied. Thus, any detailed work to examine or optimize an mHQ approach should be carried out only after chemical action lists and both AL1 and AL2 values have been selected for a program. The metal consensus ALs used in the present paper are based upon 7 to 8 disposal-at-sea ALs from around the world. With the exception of tPAH and tPCB, the organic consensus ALs, on the other hand, are based on far fewer values, and often, on values not generated for DaS programs. Thus, although they can be used here to illustrate differences in approaches, the values that should be used in a DaS program have yet to be determined.

Differences in the performance of mHQ decision rules for additional metals versus additional organics suggested that the ALs applied in the present study might not be conservative enough for some organics; this may explain why Apitz and Agius (2013) observed that the addition of many organics to the chemical action lists had a much less dramatic effect on Tier 1 chemical outcomes than did the addition of metals. For many organic compounds, including dioxins and furans, there are very few ALs related to dredging. Given this situation, the organic consensus values used in the present work were based on ecological risk assessment [ERA] rather than on dredging values, and were derived differently and intended for a different purpose.

The effects of applying an mHQ, rather than a one out/all out rule, in the Tier 1 assessment were promising, although not completely straightforward. Further work and consideration to identify the best approaches to the use of mHQs in a tiered framework is recommended. Incorporating the latest dredging-specific organic ALs and approaches to regulating organic levels in sediment may identify potential DM framework improvements beyond those identified as a result of the present work.

Addition of a Tier 1 screening bioassay

A screening bioassessment alone did not improve Tier 1 decisions for acutely toxic sediments compared with status quo chemistry alone (although it did so for sublethally toxic sediments); the best decision outcomes were observed using the screening bioassessment and an expanded chemical action list together. Specifically, the addition of a screening bioassay in the form of a solvent-based Microtox™ assay with a conservative pass criterion substantially reduced the rates of both acutely and sublethally toxic false negatives; its use, along with an expanded chemical action list reduced the number of acutely toxic sediments that were missed in Tier 1 (i.e., Class III sediments) to very low levels.

The screening bioassay, with the addition of constituents to the chemical list, consistently improved the performance, decreased Class II and III sediments, but with a cost of higher false positive (Class IV) rates; these increased 5% to 7% compared to the same chemical lists without the screening bioassessment. Thus, it seems that regulatory performance would be relatively unchanged if the current chemical action list were replaced with this screening bioassay in Tier 1, but it would be substantially enhanced by adding a screening bioassay to the current chemical parameters tested in the first tier, and progressively improved with each extended chemical action list. The true positive rates for DaS alone (Scenario 2 in Figure 2), Microtox™ alone in Tier 1 (Scenario 10 in Figure 3), and their combination (Scenario 11 in Figure 3) suggest that chemistry and the screening bioassay capture different groupings of sediments; this is much clearer when one compares Class V (sublethally toxic true positives) to the smaller population of Class VI sediments.

It should be noted that acutely toxic sediments (those that failed the acute bioassay and/or both sublethal assays) make up only 5% of the samples in the database (55 out of 1081). This is a relatively small population of samples from which to draw definitive conclusions based on small shifts in Class III and Class VI assignments (because acutely toxic samples will be classed as one or the other, depending upon the scenario). Thus, although small changes in these assignments could indicate more effective protocols, they may also be affected by very sample-specific differences that may not be generalizable.

It is possible that the screening bioassay is identifying a population of sediments with contaminants not on the chemical action lists, but it is also possible that the screening bioassay is sensitive to potential confounding factors that also are causing higher levels of bioassay failure not driven by chemical contaminants. On the other hand, it is also possible that the very conservative bioassay threshold is failing a large-enough proportion of samples (without a real link to toxicity), which coincidentally then also fail Tier 2 bioassessment, along with the large proportion of Class IV sediments, which then ultimately pass bioassessment. It is difficult to distinguish between these potential explanations with the
available data, but results do suggest that the use of a screening bioassessment in conjunction with chemistry warrants further investigation. Because a key assumption behind the present study is that Tier 2 bioassessment represents the true state, a deeper evaluation of this question is outside the scope of this work, but investigations into indicators of confounding effects is ongoing in related work.

To further consider the addition of a Microtox™ screening assessment to Tier 1, it will be necessary to examine the performance of the solid-phase photoluminescent bacteria method, which will not necessarily comprise an assessment tool that is equivalent to the solvent-based Microtox™ method evaluated in the present work. Because the solvent-based method extracts most chemical constituents from the chemical matrix (even those that may not otherwise be bioavailable), solvent-based assays are arguably more conservative, and therefore better suited to a screening level assay, than are their solid-phase counterparts. However, solid-phase tests are likely more ecologically relevant than solvent-based extractions.

In evaluating the performance of a Microtox™ toxicity test in Tier 1, the present work revealed that Microtox™ alone (Scenario 10 in Figure 3) achieved similar pass/fail rates as the current DaS chemical action list (Scenario 2 in Figure 2) at predicting Tier 2 acute toxicity. Although this is not the optimal performance for Tier 1, it may be good enough to replace chemistry in a new, CDMMMD-recommended “pre-Tier 1, screening assessment” that is distinct from Tier 1, and intended to confirm only that sediments previously subjected to an expanded chemical assessment and identified as low-risk do not warrant extensive further assessment. Given differences in the Microtox™ and photoluminescent bacteria methods used by the USEPA and EC, a careful evaluation of the performance of the solid-phase test method as a low-risk sediment screen would be needed before deciding to proceed with a Microtox-only screening assessment, but doing so could potentially represent assessment cost savings for low-risk dredging projects.

The present paper examined the implications of adding a Microtox™ test to Tier 1. It is recognized, however, that there is an opportunity for future work exploring how other biological tests would perform as screening or prescreening tools. Such work could focus on other sublethal tests in the existing Canadian DaS bioassessment battery (e.g., EC 2001, 2014), screening bioassays used to steer chemical testing in other regulatory programs, cell-based bioassays capable of detecting genotoxicity or endocrine disruption, or genomics methods being developed for toxicity identification and evaluation (TIE) purposes (Bay et al. 2012; Bay 2014).

The need for tiers

The analyses reported here suggest that an increase in the number of chemical parameters results in a moderate increase in the number of samples requiring Tier 2 bioassessment, resulting in an increased rate of correct environmental decisions. Following from this, one could argue that it may be more time and cost effective to carry out a full suite of bioassessments in parallel with chemical analyses at the screening stage. The data in the present study do not fully support such a conclusion for 2 reasons. First, most scenarios still had 20% to 40% of samples that passed in Tier 1 and would not need further assessment. If a full suite of bioassessments was required in Tier 1, these Class I sediments would be subject to bioassessment with no change to their rate of correct environmental decisions, while the rate of misassignment of the Class II and III sediments in some scenarios was quite low. Secondly, due to the perceived complexity and uncertainty of Tier 2 biological assessments, potential DaS permit applicants in Canada sometimes withdraw their applications when Tier 1 fails, choosing either not to dredge (potentially inhibiting development) or to go directly to land-based disposal, which falls into a different regulatory framework but which may or may not have fewer overall ecological and economic impacts.

Ultimately, the balance between the objective of minimizing false negatives and minimizing undue analytical burdens on applicants is a policy decision. It is not clear to what extent the higher numbers of Tier 2 outcomes will affect the decisions and behavior of applicants, but given past trends, it is not unreasonable to assume that any increase in the number of applicants required to conduct bioassays could result in a decrease in applications for disposal-at-sea permits. Such a decrease would likely result in proponents pursuing alternative and potentially less ecologically desirable disposal options (e.g., land based). Requiring bioassessment only when indicated based upon chemical results may maintain current rates of DaS permit applications. Still, some of the better performing scenarios evaluated here result in significant increases of samples requiring Tier 2 assessment (although in those scenarios, a large proportion of samples would eventually be permitted for DaS). Thus, the relatively high rate of DM found suitable for DaS following bioassessment in the present study may encourage proponents who fail Tier 1 to pursue bioassessment when they may currently do otherwise.

The need for Canadian data

The present work assumes that the sediments analyzed in this US-based database are representative of what might be encountered in the Canadian DaS program. To assist with any updates to Canada’s sediment characterization process, future efforts will be made to integrate into the data set as much Canadian data as possible. Efforts are also underway to centralize a repository of DM characterization data, including data from a range of “clean” and “contaminated” sites, so that it will be available for future data-mining studies such as this one.

Acknowledgment—This work was funded by Environment Canada, Marine Protection Programs. The Coastal and Oceanographic Assessment, Status and Trends (COAST) Branch, part of NOAA’s National Centers for Coastal Ocean Science in the Center for Coastal Monitoring and Assessment (CCMA), is gratefully acknowledged for making its extensive data sets available online. We thank Oren Perez and his
associates for their support in resolving questions on the data sets, and Lorraine Brown Read for her thoughtfull review and suggestions for future analyses.

**Disclaimer**—This paper does not necessarily represent the views of Environment Canada or any affiliations represented by the authors. References to brand names and trademarks in this document are for information only and do not constitute endorsements by Environment Canada or the authors. The authors do not intend to suggest conclusions on the potential ecological risk or regulatory status of the sediments from which the database was drawn; these samples were not collected for the assessment of ocean disposal and this review represents an analysis of only a small fraction of the data available. These data are used only to provide a data set that might realistically represent the range of sediment types that might be encountered by the Canadian DaS program, and to evaluate the potential performance of a range of DM DaS decision rules.

**Data Accessibility**—Data are from NOAA Status and Trends (NS&T) and Mussel Watch datasets recently placed online (National Status and Trends Program, http://ccma.nos.noaa.gov/about/coast/nsandt/download.aspx).

**SUPPLEMENTAL DATA**

Supplemental file provides scenario details and discusses the issue of potential confounding factors.

**Figure S1–2.** Percentage outcomes for scenarios with various Tier 2 approaches. Scenario labels, numbers, conditions and outcomes as described in Tables 5 and 7. Classifications as in Figure 1 (paper).

**Table S1-1.** ALs used in the present study

**Table S1-2a.** Scenario conditions, Scenarios 1 to 14

**Table S1-2b.** Scenario conditions, Scenarios 15 to 27

**Table S1-2c.** Scenario conditions, Scenarios 28 to 40

**Table S1-2d.** Scenario conditions, Scenarios 41 to 47

**Table S1-2e.** Scenario conditions, Scenarios 48 to 54

**Table S1-3a.** Scenario outcomes, Scenarios 1 to 14

**Table S1-3b.** Scenario outcomes, Scenarios 15 to 27

**Table S1-3c.** Scenario outcomes, Scenarios 28 to 40

**Table S1-3d.** Scenario outcomes, Scenarios 41 to 54

**REFERENCES**

Agius SJ, Porebski L. 2008. Towards the assessment and management of contaminated dredged materials. *Integr Environ Assess Manag* 4:255–260.

Apitz SE. 2008. The derivation and application of sediment quality guidelines (SQGs) and bioaccumulation models in the screening and first tier of the assessment of sediment quality. Report for the Canadian Disposal at Sea Program. Little Hadham (UK): SEA Environmental Decisions. 90 p.

Apitz SE. 2010. Review of the use of bioassays for Canada’s Disposal at Sea Programme - Final Report to Environment Canada. Little Hadham (UK): SEA Environmental Decisions. p 180.

Apitz SE. 2011. Integrated risk assessments for the management of contaminated sediments in estuaries and coastal systems. In: Wolanski E, McLusky DS, editors. Treatise on estuarine and coastal science. Waltham (MA): Academic. p 311–338.

Apitz SE, Agius S. 2013. Anatomy of a decision: Potential regulatory outcomes from changes to chemistry protocols in the Canadian Disposal at Sea Program. *Mar Pollut Bull* 69:76–90.

Apitz SE, Barbanti A, Giulio Bernstein A, Bocci M, Delaney E, Montobbio L. 2007. The assessment of sediment screening risk in Venice Lagoon and other coastal areas using international sediment quality guidelines. *J Soils Sediments* 7(5):326–341.

[ASTM] American Society of Testing and Materials. 1992. Standard guide for conducting 10 day static sediment toxicity tests with marine and estuarine amphipods. In: Annual Book of ASTM Standards, Section 11, Water and environmental technology. Philadelphia (PA): E 1367-90. p 1083–1115.

Bay S. 2014. Molecular tools for toxicity identification evaluation. Costa Mesa (CA): Southern California Coastal Water Research Project. [cited 2015 Feb 13]. http://www.sccwrp.org/ResearchAreas/Contaminants/MolecularToolsForToxicityIdentificationEvaluation.aspx

Bay S, Vulpe C, Loguinov AV, Araì A. 2012. Application of gene expression analysis for sediment toxicity stressor evaluation. Costa Mesa (CA): Southern California Coastal Water Research Project. Technical Report 660.

Brack W, Klammer HJC, Alda MLd, Barceló D. 2007. Effect-directed analysis of key toxicants in European river basins: A review. *Environ Sci Pollut Res* 14(1):30–38.

Carr RS. 1998. Sediment porewater testing. In: Clesceri LS, Greenberg AE, Eaton AD, editors. Standard methods for the examination of water and wastewater, Section 8080, 20th ed. Washington (DC): American Public Health Assn.

Daughton CG. 2002. Pharmaceuticals and personal care products (PPCPs) as environmental pollutants – Pollution from personal action. Las Vegas (NV): Environmental Sciences Division, National Exposure Research Laboratory, Office of Research and Development, USEPA.

[EC] Environment Canada. 1998. Biological test method: Reference method for determining acute lethality of sediment to marine or estuarine amphipods. Ottawa (ON): Method Development and Applications Section, Environment Technology Centre. Reference method EPS 1/RM/35.

[EC] Environment Canada. 2001. Biological test method: Test for survival and growth in sediment using spionid polychaete worms (Polydora cornuta). Ottawa (ON): Method Development and Applications Section, Environment Technology Centre. EPS 1/RM/41.

[EC] Environment Canada. 2002. Biological test method: Reference method for determining the toxicity of sediment using luminescent bacteria in a solid-phase test. Ottawa (ON): Method Development and Applications Section, Environment Technology Centre. Reference Method EPS 1/RM/42.

[EC] Environment Canada. 2011. Biological test method: Fertilization assays using echinoids. 2nd ed. Ottawa (ON): Method Development and Applications Section, Environment Technology Centre. Report EPS 1/RM/27.

[EC] Environment Canada. 2014. Reference method for measuring the toxicity of contaminated sediment to embryos and larvae of echinoids (sea urchins and sand dollars). Ottawa (ON): Method Development and Applications Section, Environment Technology Centre. Report 1/RM/S8.

[IMO] International Maritime Organization. 2009. Guidance for the development of action lists and action levels for dredged material. London Convention, London Protocol. London (UK) p 36.

Lahr J, Maas-Diepeveen JL, Stuijfzand SC, Leonards PEG, Druke JM, Lucker S, Espeldoon A, Kerkum LCM, van Stee LLP, Hendriks AJ. 2003. Responses in sediment bioassays used in the Netherlands: Can observed toxicity be explained by routinely monitored priority pollutants? *Water Res* 37:1691–1710.

Long ER, Hameedi J, Robertson A, Dutch M, Aasen S, Ricci C, Welch K, Kammin W, Carr RS, Johnson T et al. 1999. Sediment quality in Puget Sound: Year 1 - Northern Puget Sound. Silver Spring (MD): US Dept of Commerce National Oceanic and Atmospheric Adm, National Ocean Service, National Centers for Coastal Ocean Science, Ctr for Coastal Monitoring and Assessment. Report nr NOS NCCOS CCMA No. 139, Technical Memo No. 139. Jointly published by Olympia (WA): Washington State Department of Ecology Environmental Assessment Program, Environmental Monitoring and Trends Section. Publication No. 99-347, 350 p.

[USEPA] US Environmental Protection Agency. 1993. Guidance manual: Bedded sediment bioaccumulation tests. Washington (DC). EPA/600/R-93/183.
Wagner WE, 1995. The science charade in toxic risk regulation. Columbia Law Rev 95:1613–1723.
Webster L, Roose P, Bersuder B, Kotterman M, Haarich M, Vorkamp K. 2013. Determination of polychlorinated biphenyls (PCBs) in sediment and biota. Copenhagen (DE): International Council for the Exploration of the Sea. p 18. No. 53, ICES Techniques In Marine Environmental Sciences.
Wenning RJ, Batley GE, Ingersoll CG, Moore DW. 2005. Use of sediment quality guidelines and related tools for the assessment of contaminated sediments (SQG). Pensacola (FL): SETAC. p 783.