MTCA Science Panel
May 20, 2011
Meeting Summary

Location

University of Washington, Botanical Gardens, Seattle, WA

MTCA Science Panel Members present

Dr. Bruce Duncan
Dr. Elaine Faustman
Dr. Teri Floyd
Dr. Mike Riley
Dr. Rosalind Schoof

Ecology staff

Chance Asher
Dave Bradley
Martha Hankins
Laura Inouye
Craig McCormack

Audience members

Marion Wineman, WR Consulting
Nicole Rangel, URS Corp
Burt Shephard, US EPA
Lon Kissinger, US EPA

Presentations

Russ McMillan (Ecology)
Teresa Michelsen (Avocet)

Meeting Summary

The meeting started at 9:10 A.M. Martha Hankins welcomed everyone and reviewed the agenda. Members submitted a couple of minor corrections to the August 25, 2010 meeting summary. Dr. Duncan provided opening comments and mentioned that the peer review conducted by

Ecology of the freshwater sediment standards employed international experts. He noted that considerable freshwater experience is in the room here, and introduced Burt Shephard.

General Overview of Updates to the MTCA and SMS Rules: Dave Bradley

Dave provided an overview of Executive Order 10-06 halting non-essential rulemaking. He provided an overview of the range of issues the program is currently working on and expects to address in the next few years. He noted that Ecology is working on TEE guidance. There are a range of things Ecology is thinking about, including a look at compounding conservatism; will want to look at this moving forward.

Dave reviewed how freshwater sediment criteria fits within the current SMS framework. He reminded members that at the last Science Panel meeting Ecology described the biological and chemical criteria being developed for managing risk from contaminated sediments using the existing two-tiered framework. He noted both scientific and policy issues. Scientific issues include verifying the choice of appropriate species and bioassays for characterizing freshwater benthic communities. Policy choices include decisions around risk management, level of protection, and the balance between false positive (where clean sites are incorrectly identified as contaminated) and false negative (where contaminated sites are missed) rates.

He noted the scientific approach taken by the Ecology Sediment Group evolves and absorbs new science as it becomes available.

Presentation: Russ McMillan (Freshwater Biological Standards)

Russ briefly reviewed the material presented at the last meeting. He noted that the existing regulatory framework for decisions around contaminated sediments consists of two tiers: a lower no-adverse effects level - the Sediment Quality Standard (SQS) - and a higher minor adverse effects level - the Cleanup Screening Level (CSL).

The CSL serves as an upper bound for allowable sediment concentrations; the SQS is set based on minimal detectable differences between the sample and a control or reference.

Discussion

Dr. Faustman asked for clarification of Appendix A and the proportion of chemicals at a site and how certain analytes were treated.

Russ described the work looking at the proposed suite of bioassays, the focus on the function of the benthic community in the ecosystem, and the species proposed for the bioassays as an indicator of benthic community health. So instead of just looking at whether the critters were present, we are asking whether the function they provide is preserved.

The two species proposed, the midge *Hyalella* and the amphipod *Chironomus* display different life strategies and are key to the function of freshwater benthic communities here in the Northwest. The amphipods are shredder grazers, while midge larvae hatch in the sediments and as they wiggle toward the surface, provide a primary food source for animals such as migrating ducks. Both species are instrumental in getting organic material recycled back into the sediments. Benthic communities are dependent on the key functions provided by these two organisms. When these functions are compromised, organic matter accumulates and the aquatic system can become anoxic.

Members discussed that a conceptual model or flow diagram would be helpful.

Members asked about biomass as an endpoint. It was noted that Biomass is a population endpoint not an organism endpoint.

Dr. Duncan noted that *Hyelella* is a good representative of the life strategy for this particular type of organism.

Panel members noted the rationale for the choice of the two species, including their presence in Pacific Northwest ecosystems and their appropriate level of sensitivity to environmental contaminants.

Panel members asked about biological criteria endpoints. Russ clarified that the bioassays are run against control samples instead of references. The controls are laboratory standards using clean sediments. Ecology has been unable to locate adequately pristine freshwater reference areas in Washington. Panel members noted that because of a lack of freshwater reference sites in Washington, Ecology should include flexibility for site-specific use of bioassays and recommended that rule or guidance include bullets with factors to consider when employing the site-specific bioassays or chemical criteria.

Russ reviewed Ecology responses to questions asked at the August 25, 2010 Science Panel meeting. Panel members expressed appreciation for the written responses provided in the meeting materials. Members asked for additional information about how sensitive the chosen bioassays are for the chemicals typically found at contaminated freshwater sites in the Pacific Northwest. It was noted that this is important for establishing a clear link between the biological effects and the chemical signature.

Ecology confirmed that the range used for grain size and pH is consistent with a majority of lakes and streams in Washington. Examples of Moses Lake and Soap Lake were discussed, as well as the unique situation provided by the Columbia River. It was noted that grain size is the largest variable. Members suggested that a conceptual model would be useful. Ecology noted that some systems may appropriately consider other species; systems with snails may choose that critter as an appropriate for evaluating the benthic community. Asked about using a biomass endpoint (the sum total weight of surviving critters in the bioassay), Dr. Shephard explained that this is a population density effect: critters could flunk the survivability test (that is, more die) but those surviving grow really well due to a lack of competition for food. So a fewer number of critters could exhibit greater a biomass. Ecology noted that there is currently no American Society for Testing and Materials (ASTM) standard for this endpoint.

Dr. Faustman noted that the question asked of the peer reviewers “is the proposed bioassay suite appropriately sensitive to protect the freshwater macro benthic community (i.e., typical taxonomic structure and function such as prey base to endangered species such as salmon)” was slightly different from the question posed to the Science Panel because the proposed criteria are not aimed to be protective of bioaccumulative effects to higher trophic levels. She asked for clarification regarding the quality of the food base – that it’s fundamental as to whether the question is quality or quantity. Ecology responded that the criteria are protective of the benthic community at a population level, not individuals or species levels, and was not intended to address bioaccumulatives. Ecology has a proposed narrative for ecological bioaccumulative effects.

Presentation: Teresa Michelsen – chemical criteria

Teresa reviewed the questions presented to the Science Panel regarding the method being used to develop freshwater Sediment Quality Guidelines (synonymously called Sediment Quality Values [SQVs]). The floating percentile method is being used to develop chemical criteria at two levels: the lower Sediment Quality Standards (SQS) and the upper Cleanup Screening Level (CSL). She noted the regulatory framework (having the two tiers) is consistent between marine and freshwater sediment standards, and explained the reason for using a different method to develop freshwater sediment standards.

Question for Discussion: Is it appropriate to use sediment bioassays to represent effects to the benthic community? Is the use of a multivariate model to empirically derive chemical SQVs scientifically defensible?

Members discussed the differences between the marine and freshwater methods and environments, and noted that freshwater environmental conditions make a multivariate model attractive because of the diverse nature of freshwater systems.

Ecology noted that the difference between the floating percentile method and other methods used to develop freshwater standards is that this method is multivariate and incorporates the relationship between multiple contaminants. Being an empirical model, the test of success is the ability to predict observations.

Question for Discussion: Is the data set sufficiently robust and representative? Has appropriate data screening and quality assurance been conducted?

Teresa noted that Ecology is confident in the data set developed.

Question for Discussion: Is the reliability testing that was conducted an appropriate method for evaluating SQVs? Is the comparative reliability analysis that was conducted an appropriate way of making decisions while fine-tuning the approach?

SQVs developed using the Floating Percentile Method (FPM) and a robust data set were subjected to reliability testing.

Science Panel members asked about prevalence of adverse effects. Ecology responded that measures of reliability are indicative of prevalence. The RSET team has looked closely at comparative reliability: they analyzed data from east and west parts of the region; they looked at chemicals (for example, TPH verses PAHs); comparisons to the *Microtox* test; and only *Hyalella*.) Ecology noted that an appendix is being added to the report showing RSET reliability test results.

Because of variability in the freshwater environments, it was important to use the entire dataset in developing the SQVs.

Members asked about prevalence in the dataset.

The data set includes all paired chemical and biological data for sediments in Oregon, Washington and Idaho. This made validation complicated. Dr. Schoof noted that it might be useful to inquire about data in British Columbia.

There was considerable discussion about the reliability statistics. Ecology noted that the SQS is intended to be conservative; the CSL on the other hand is the upper bound and allows some adverse effects to the benthic community. The difference between risk management and risk assessment was noted.

Following the peer review an additional statistical analysis (Bias, Odds Ratio, Hanssen-Kuipers Discriminant, Post-Test Odds Ratio) was performed as recommended by EPA specifically to address questions of prevalence and reliability. (That is, because the data included both hit and no-hit data, how did this affect the results? Given the presence of the no-hit data in the dataset, was a 20% false positive rate sufficiently conservative?) These different reliability tests were run for each biological endpoint and for the combined endpoints.

Bias (the ratio of the samples predicted to be toxic divided by the samples that are actually toxic) provides another measure of reliability and is a measure of conservatism. The goal is for the bias to be equal to or greater than 1; this way, the SQVs err on the side of being appropriately predictive of toxicity or slightly over predictive. The analysis indicated for the SQS a bias of 1.7 and for the CSL a bias of 1.1.

The odds ratio is calculated as the likelihood that a prediction that a sample is toxic or nontoxic is correct over the likelihood that the prediction is incorrect. The odds ratio of 4:1 for the SQVs suggest an 80% likelihood of exceeding the biological standards given a chemical SQV exceedance.

Panel members asked Ecology to back up and explain some of the assumptions. Ecology noted that the problem of developing freshwater standards has been broken in multiple steps: (1) identify appropriate biological criteria (2); develop a robust data set of paired chemical and biological data; and, (3) use the FPM applied to the data set to establish acceptable SQVs.

The goal is to establish which chemicals and concentrations of those chemicals best predict toxicity to the benthic community as measured by bioassay results. The FPM set the SQS and CSL for each chemical by predicting the outcome of the bioassays based on empirical data.

Audience Comment

Dr. Shephard noted that much of this type of analysis had been performed in Portland Harbor, and information is available on the Sediment Management Annual Review Meeting (SMARM) website.

Working lunch

Panel Members asked for a table that indicates the number of chemicals observed for each sample. Ecology noted that this is huge. Members clarified they are asking about what chemicals were observed and the range of concentrations.

Ecology described the challenge of setting the SQS and CSLs. The goal was to develop the two levels for use in regulatory decision-making. The results of the analysis were a bit surprising in that the species differences were greater than the endpoint differences. That is, chemical concentrations predicting the growth and mortality results for each species were close together. So concentrations generally affecting each species were separated more than concentrations affecting either growth or mortality in general. Consequently, the values of the bioassay

endpoints were combined into a single distribution. The SQS was selected as the lowest value and the CSL was selected as the next highest significantly different value, to provide conservative values at the low end of the no-adverse-effects to minor-adverse-effects distribution.

Ecology noted use of the EIM database. Members discussed table 3-7, co-location, uncertainty, types of field variability, and laboratory procedures. It was noted that although analytic chemistry for sediments is a crude indicator of toxicity, it could provide an expedient way of screening sites.

Science Panel members offered to provide information from their experiences on sources of variability in sediment chemistry. Ecology noted that this is an empirical model. As, Cr, V, and Sb were discussed.

Science Panel members generally expressed support for the approach and analysis; they would like some additional information about the various chemicals encountered in data set used in developing the SQVs.

Summary of Science Panel Responses to Discussion Questions

Is it appropriate to use sediment bioassays to represent effects to the benthic community?

- Members agreed that the species and endpoints chosen represent the best available science. It was noted that the question presented to the panel is more general than what was discussed. Members noted the need for regulatory flexibility.

Is the use of a multivariate model to empirically derive chemical SQVs scientifically defensible?

- Members agreed that the use of a multivariate model to empirically derive chemical SQVs is within the range of scientific defensibility, but noted some qualifications. Multivariate analysis has limitations. Causality is not identified, and correlation doesn't imply causation.
- Limitations for wetlands were noted. Ecology acknowledged that the dataset does not include wetlands and noted that wetlands might be an appropriate topic for another meeting.

Is the data sufficiency/reliability testing for the dataset used in the floating percentile method scientifically defensible?

- Members noted that the dataset is very well screened and that Ecology did a very good job in quality assurance.
- A flow diagram was recommended to show how the reliability testing is used and how new data could be used to update the criteria.
- Members noted the importance of validating data and acknowledged the role of the Sediment Management Annual Review Meeting (SMARM) in applying new science as it becomes available.
- It was noted that multivariate statistical approaches were developed to predict unlikely and rare events. These methodologies come from the fields of meteorology and epidemiology, and are used in the study of rare diseases.

Validation: Is the overall framework and selection of final SQVs scientifically defensible?

- Panel members agreed that the floating percentile method used to develop SQVs is a robust method. They noted that the scientific defensibility is linked to the application and the flexibility around implementation

Use of California EPA Toxicity Values to Establish MTCA Cleanup Levels: Dave Bradley

Ecology is faced with how to establish toxicity for frequently encountered substances without toxicity values in EPA's Integrated Risk Information System (IRIS) or with out-of-date values in the Health Effects Assessment Summary Tables (HEAST). California EPA (CalEPA) has developed toxicity values for a wide range of substances using policies and procedures similar to EPA. Ecology is looking for a way to programmatically include recent scientific data, and notes that HEAST has not been updated since 1997.

MTCA establishes a hierarchy for use of toxicity data:

1. Integrated Risk Information System (IRIS)
2. Health Effects Assessment Summary Tables (HEAST) or if more appropriate the National Center for Environmental Assessment (NCEA)
3. Other values established using methods described in EPA Guidance (non-carcinogens) or in the MTCA rule or with other methods if there is clear and convincing scientific data (carcinogens); in consultation with the MTCA Science Advisory Board, Department of Health, and EPA, and according to the MTCA procedures for considering new scientific information in WAC 173-340-702(14), (15), and (16).

Ecology is proposing that cancer slope factors and inhalation unit risk factors from other sources may be used to establish cleanup levels and remediation levels when values are not available in the IRIS database. The Department is proposing to use the criteria in the EPA Office of Solid Waste and Emergency Response (OSWER) Directive 9285.7-53 when evaluating whether particular values can be used to support decisions on cleanup levels or remediation levels.

Ecology is considering the following approach to decide whether to use individual California EPA Toxicity Values to Establish MTCA Cleanup Levels:

- Screen available California EPA values and place substances/toxicity values into one of 3 review categories (high, medium, or low priority)
- Perform reviews for priority substances using the quality of information criteria specified in the MTCA rule, WAC 173-340-702(16)
- Consult with the Science Panel, Department of Health, and EPA as appropriate
- Provide opportunity for public comment
- Prepare final decision and update the Cleanup Level and Risk Calculation (CLARC) database
- Provide a clear statement in site-specific documents when CalEPA values are used to establish cleanup levels for individual sites

Audience comment

Lon Kissinger (EPA Region 10 offered to contact Dave Crawford regarding use of Regional Screening Level (RSL) values as a source of toxicity data.

Discussion

Members noted that the Provisional Peer Reviewed Toxicity Value (PPRTV) tables from EPA are now publically available. They asked Ecology to clarify the details of the public review used by CalEPA, the Agency for Toxic Substance and Disease Registry (ATSDR), and for the PPRTVs.

Members recommended that Ecology identify the type of review process each organization uses.

Members discussed various chemicals (including TCE and naphthalene). They noted that Ecology should not automatically adopt CalEPA values, but that the review process proposed by Ecology is appropriate. They noted that Ecology's Cleanup Level and Risk Calculation (CLARC) database is a useful tool.

The Panel approved the August 25, 2010 meeting summary.

Meeting summary approved on June 11, 2012