

**Sediment Workgroup  
May 3, 2010 Meeting Summary**

**Location**

Tacoma Wastewater Treatment Plant, Tacoma, WA

**Sediment Workgroup Members Present**

Joanne Snarski, Paul Fuglevand (not attending), Lon Kissinger, Glen St. Amant, Teresa Michelsen, Pete Rude (attending by phone), Clay Patmont, Jack Word

**Ecology Participants**

Chance Asher, Dave Bradley, Laura Inouye, Fu-Shin Lee, Russ McMillan, Pete Adolphson, Dave Sternberg, Donna Podger, Stacie Singleton.

**Audience**

Lawrence McCrone, Rob Pastorok, Roger McGinnis, Jessi Belston, Todd Thornburg, Heather Trim, Denice Taylor

**Agenda**

1. Freshwater Standards SQV Technical Report Discussion
2. Human Health/Background Discussion
3. MTCA/SMS Integration: SMS Definitions, Harmonization of Terminology

Meeting notes are organized by the agenda number and include a short summary of material presented with a brief summary of the main discussion points. Detailed notes on the discussion are included in an appendix.

**Acronyms**

AET – Apparent Effects Threshold

ASTM – American Society for Testing and Materials

BPJ – Best Professional Judgment

CERCLA – Comprehensive Environmental Recovery, Compensation and Liability Act

CSL – Cleanup Screening Level

COC – Chemical of Concern

EPA – US Environmental Protection Agency

EPH – Extractable Petroleum Hydrocarbon

ESA – Endangered Species Act

FW – Freshwater

IBI – Index of Biotic Integrity

LDW – Lower Duwamish Waterway

MCL – Minimum Cleanup Level

MTCA – Model Toxics Control Act

NOAA – National Oceanic and Atmospheric Agency

ODEQ – Oregon Department of Environmental Quality

PAH – Polycyclic Aromatic Hydrocarbon

PCB – Polychlorinated Biphenyls

PLP – Potentially Liable Party

QA – Quality Assurance

QC – Quality Control  
RSET – Regional Sediment Evaluation Team  
SAPA – Sampling and Analysis Plan Appendix  
SEM-AVS – Simultaneously extracted metal – acid volatile sulfide  
SMS – Sediment Management Standards  
SQG – Sediment Quality Guideline  
SQS – Sediment Quality Standards  
SQV – Sediment Quality Value  
TMDL – Total Maximum Daily Load  
TPH – Total Petroleum Hydrocarbon  
TOC – Total Organic Carbon  
USEPA – United States Environmental Protection Agency  
VPH – Volatile Petroleum Hydrocarbon  
WQ – Water Quality

## Meeting Summary

### 1. Freshwater Standards SQV Technical Report Discussion

#### a. Freshwater Issues Update/Presentation – Russ McMillan

- Freshwater issues will include biological and freshwater criteria. Looking at how to select among the suite of bioassays available. We'll discuss next steps for Ecology for rule revision adoption.
- SQV Report is being peer reviewed by both this group and the Oregon DEQ. Need comments back by May 17. Ecology will then present it to the Science Panel, find other peer reviewers, including the EPA.
- Ecology is looking at adopting both chemical and biological standards for freshwater.
- Ecology's preference for biological standards is to use a minimum of 3 species across 3 bioassays for best sensitivity representation. Looking at which endpoints to use.
- Have enough data to calculate guidelines for both acute and chronic toxicity. Ended up not using Microtox. Want to use the *Chironomus* chronic test as part of the suite even though not quite enough data for guidelines yet. Not the first time we've done this.
- One option would be 3 distinct endpoints in 2 or more tests. But we need to be careful of how we use the endpoints.
- The freshwater will have a consistent framework as the current SMS marine standards.
- Exposure duration discussion around 10- and 28-day *Hyaella* test; what durations are acute and chronic. Caution to keep terminology straight.

## **Sediment Quality Values Report Discussion/Presentation – Teresa Michelsen**

### **• Presentation:**

- The Sediment Quality Value (SQV) Report originally came out of a RSET workgroup to develop freshwater guidelines for Washington, Oregon, and Idaho for both dredging and cleanup. This method was developed to address the shortcomings of other methods.
- The method tries to minimize false negative and false positives. Can identify classes of metals and of organics affecting reliability.
- Started with a dataset of 300 chemicals. Screened data at a high level of QA/QC, useable data. Fairly automated. Overlap between red and green areas implies the area is sometimes toxic and sometimes not.
- Predicted hit and predicted no-hit reliability influence the regulated community's perception of whether or not the SQVs are accurate.
- Substantially increased dataset from 2003, but would still like more data. With the data set, we tested a large range of differences from controls and picked the ones with the strongest correlation between chemistry and bioassays. Typical reliability results were presented and discussed. The next step is to select criteria then go through peer review. Big challenge to look at how to apply the standards in large geographical areas. Also a challenge to deal with false positives and reference area QA/QC. Ecology may not want to put everything in rule, to provide some flexibility.

### **Discussion:**

- It's been a challenge to address what we should be protecting and determining if no toxicity exists as opposed to setting the lower range at an acceptable level of false negatives.
- It is critical for the freshwater guidelines to expect these sites to be very different from one another, much more so than marine sites. Freshwater sites have increased variability.
- Ecology is looking at including language in the rule for precautions to assess if a site is likely to be representative of the data set the criteria were developed from. A recommendation was made to add more guidance to the SAPA.
- RSET looking into how organisms are impacted from bioaccumulatives and toxicity. Modeling can show these impacts and the policy issue is the balance of no impact to more severe impacts.
- East versus west data set and reliability was discussed. Overall reliability was increased using the dataset as a whole.
- Regarding TPH and diesel ranges and organic carbon, what is the best assessment in relation to bioavailability? RSET looked at some of these issues and found that using TPH is best, on a dry weight basis.

- Metals assessment might be greatly improved if we look at pore water. If anyone has a good indication of where SEM and AVS plays a good role in developing criteria, please let Ecology know. However some with experience in SEM and AVS didn't find it very helpful.
- It's not clear that we can predict any better than we do with bulk sediment measurements. An attempt to correlate TPH with PAHs didn't work. The environment just isn't as predictable as partitioning theory suggests.
- We should look at how the current bioassays or other species track with or correlate with the species out in the field that they are supposed to be surrogates for. And how they track with community level endpoints. We did look for freshwater benthic community data and couldn't find any.
- Discussion of the number of samples needed in site decisions. Site managers are instructed to use a minimum of three samples for site listing decisions.

## **2. Human Health and Background Summary Document Discussion**

- Ecology is not expecting group consensus recommendations. Ecology reviewed all feedback from the advisory groups on these two issues and developed a conceptual framework to reflect this feedback and some of Ecology's original thinking. We are responding to what we heard from this group about what Ecology is doing with the feedback we've heard. Ecology appreciates the great feedback.
- Regional background definition is challenging and we will likely never get two people to agree on the same numbers. It seems to hold promise but many complex details.
- Discussion of submitted comments from the Sediment Workgroup:
  - Alternatives needed for the regional background phrase 'hydrodynamically influenced.' Need a definition that doesn't allow influences by a group of facilities.
  - BOLD data example: Could implement a cross program strategy and look at all the major urban bays for data collection. BOLD data will be relevant and useful in many cases, though.
  - Need the regional background data before implementing the rule. Need a basis for how to define regional background. Determine where 'natural background' would apply.
  - Draw boundaries within which regional background applies.
  - Recontamination and storm water sources are complex. Develop a recontamination potential standard? Recontamination is a societal issue. A lot of cleanup is from legacy contamination. The new contamination or recontamination is from stormwater and the source control component needs to be addressed on a larger scale. Are Ports supportive of our discussions about providing final settlements for areas that can be defined by chemical signature or remaining concentrations? Yes, generally. Recontamination is often

expected. Working with the PLPs to make it all come together and move forward is good. How do we move forward and take a phased approach?

- Cleanup should focus on legacy areas and source control should focus on the remaining reductions in contamination over large areas.
- How should Ecology deal with long term goals for an entire embayment to get back down to natural background? Is this a cleanup issue?
- Stormwater and atmospheric input – MTCA may not be the right tool for these. Be careful with language in the rule.
- TMDL programs - do we integrate or work with these programs? Recently TMDL programs have taken sediments into account as a potential source.
- What source control do we need to produce the results (reduced contaminant concentrations) we want? Look at both bioaccumulative risks as well as direct contact risk.
- What about risk based levels for human health, going to lead to below background levels? Is there any parallel evaluation of risk assessment methods for bioaccumulation based values? Ecology isn't going to do that right now but is looking at fish consumption rates.
- Look at RSET's data for biological receptors. Also, there are several levels for human health consumption.
- Look at basis for the standard and how you measure compliance.
- Elliott Bay is difficult for assessment. Treated sewage discharged into it as well as the industrial discharges are major complications that need to be taken into account. Take into account sources when determining regional background.

### **3. MTCA/SMS Integration: SMS Definitions**

- Trying to integrate terms found in MTCA and in SMS to clarify and make simpler. Terms being addressed include:
  - RI/FS integration on alternatives
  - Chronic
  - Sediment (never actually defined)
  - Active Cleanup Action
  - Biologically Active Zone

## **Appendix: Detailed Notes of May 3, 2010 Sediment Workgroup Meeting**

CHANCE ASHER I'm going to open it up to Russ McMillan and Teresa Michelsen and let them run the show for the freshwater piece. Just to let you guys know, and remind you what we're looking for in terms of feedback is for you to look at this as a scientific peer review process. We also would like feedback from a bigger picture policy perspective. This is also not the only time you can provide feedback on the freshwater piece but this is an opportunity to improve the written technical report in the next couple weeks. I'll turn it over to Russ McMillan and Teresa Michelsen.

RUSS MCMILLAN Just want to give you just a quick update on some of the issues with freshwater. Teresa will spend more time focusing on the completion of the sediment quality value (SQV) report. What we want to get out of the group today is comments on the development of the SQVs, if you've had a chance to review the report. We want to open up discussion and answer any clarification questions that may help continued review. Spend a little amount of time today on the biological criteria framework. We are going to include bioassays as part of freshwater criteria. And we want to look at how we would select amongst the suite of bioassays we have available. Touch on some of the next steps that are coming up for Ecology in terms of getting rule revisions adopted and look at those proposals. So in previous meeting December of last year and January of this year we caught up with where RSET was in developing sediment quality values. The effort under RSET had come to a loggerhead over a couple of policy/technical issues. About late November, early December, we began to see movement between technical staff on both sides of the border – Washington and Oregon - and to get out of some of the conundrums. One of the issues discussed at that meeting (helped to open floodgates and get some things moving) was the proposed bioassay suite. Identifying some other endpoints helped resolve some of the stalemate regarding developing numeric criteria. One of the things considered just briefly at the December meeting, was Teresa taking one last look at our data set, pulling those data out of that dataset with a new filter using different bioassay interpretive guidelines. She had done that and will touch on details with the selection of bioassay endpoints. With input from this group and two different labs we worked with Newfields and a lab in Newport (NAS), and a national expert, Chris Ingersoll.

So where are we at right now? We are looking at scientific peer review of the SQV report, this group being one of those levels of peer review. By the way, comments are due by May 17. Oregon DEQ also has the document and has provided a review. One of the next steps is to take it before the Science Panel (was MTCA Science Advisory Board). They review the technical and science issues associated with MTCA and SMS revisions. EPA will be also combing through the development of the criteria because the SMS are considered water quality standards and EPA oversees that. So when we make changes to freshwater criteria, whether it is sediment standards or water quality standards, those are subject to EPA approval.

When we first came out of the gates on developing freshwater sediment criteria, we were looking at probably just adopting biological criteria for our rule revision. We felt there was still a lot to be sorted out for the numeric chemical criteria. Those have since come together rather quickly with regards to our timeline. We are now looking move towards adopting both chemical and biological standards. The Program Management Team within TCP has approved inclusion of

chemical criteria and biological criteria for further peer review. One of things I'm looking forward to today is focusing some time in this group on the biological criteria and framework. The existing marine standards work under assumption of requiring three bioassays and those are distributed as two acute and one chronic test. The preference is to employ a minimum of 3 species across those 3 bioassays to get the best representation of the sensitivities that might be exhibited by the benthic community. It's a little problematic to be able to require those same details given a handful of bioassays that we have endpoints for and would look to adopt as part of freshwater standards.

We have an array of bioassays (chronic and acute) and endpoints (lethal and sub-lethal) to choose from. Then we have Microtox, where there is always a bit of a question whether that should be considered a chronic test or not? If it is considered a chronic test, Ecology would have to determine whether to allow using Microtox as the only chronic test in a suite of 3 tests.

Selecting the minimum requirements for a suite of 3 bioassays must take into account the range of species and sensitive life-history stages typical of a benthic community. Options would favor multiple species and including at least one chronic test that examines population level effects or endpoints.

One of the other options we have to play with amongst our handful of bioassays are the 20-day *Chironomus* endpoints. Though not used in the development of the Freshwater SQVs, these are promising additions to the bioassays and endpoints available.

Here's the suite of bioassays we currently have before us. The oval circles represent the changes made from the previous way that endpoints were determined to be hits or no hits under the SQS and CSL. What I'd like to do here is spend some time brainstorming how we would select from amongst this suite of bioassays; the choice of 3 different bioassays, with at least 1 chronic or sublethal, and preference for 3 species. I'd like to toss it out for you guys to take a look at what kinds of options we should consider.

TERESA MICHELSEN I have one comment. I think on a previous slide you meant *Hyaella* 28-day. What we have now is enough data to calculate guidelines for acute and chronic *Hyaella* test and the acute *Chironomus* test. We've had some issues with past Microtox data, and we ended up not using it. There's probably nothing wrong with current protocol but I have a lot of questions with calling it a chronic test or sub lethal test or even a freshwater test. I would like to see the *Chironomus* chronic test be available for use as part of the suite even though we didn't base our guidelines on it. One of reasons we didn't base guidelines on it was that we didn't have enough data. We have 30 data points and it is of interest to the region to have more data for that test. So it wouldn't be the first time as a program that we've encouraged the use of a test we haven't developed criteria for yet. I would like to go ahead and do that. And as far as rule revision, you could say 3 endpoints, at least one of which has to be chronic but require 2 or more species. That way we do eventually get the data, well we've looked at a lot of other freshwater species from other regions and they just don't work here. They're the wrong temperature or wrong environment. I don't know how long it's going to be before we get a 3<sup>rd</sup> species.

RUSS MCMILLAN Can we throw up some of the options on the board?

TERESA MICHELSEN We're working on the guidelines; but those are some of my thoughts.

RUSS MCMILLAN Well, I was thinking that maybe we could toy with some of these as the numbered options or something along those lines.

TERESA MICHELSEN So I keep the acute and chronic but just change it to endpoints rather than species or tests -3 distinct endpoints and 2 or more species.

RUSS MCMILLAN Call this Option 1 – it would be 3 distinct endpoints.

TERESA MICHELSEN And 2 or more tests, two or more species, and retain the acute and chronic and not get wrapped around the sub lethal.

RUSS MCMILLAN Retain at least one chronic.

TERESA MICHELSEN Right. Retain the need to have acute and chronic in the suite. If you're running an acute test and a chronic test, and one of those has 2 endpoints, sometimes it's the chronic test that has two endpoints. So you might want 2 acute and 1 chronic or 1 acute and 2 chronic. If you're running two distinct tests 2 different species, you have 3 endpoints or 4 endpoints some times.

PETE ADOLPHSON A question I have is, and this comes up often, if you have 1 test and it has 2 endpoints, do you use both of those, or do you run separately because the two endpoints will be linked to one another?

TERESA MICHELSEN They're working nationally on some changes to the endpoint definitions to try and reduce the impact of one on the other as far as calculation metrics. So I think you could go ahead and use both for efficiency's sake but be careful how you use them. For the purposes of calculating the guidelines, because they've been linked in the past, I just removed, for example, from the chronic growth test any samples that had such severe mortality that it would impact the results of the growth test that might bias.

PETE ADOLPHSON Right. So there would have to be criteria established for what severe mortality is.

TERESA MICHELSEN Right, or move to a biomass endpoint or some other endpoint that doesn't have this big of an issue. Fortunately, because I did remove those samples, these interpretive guidelines could be used equally with growth or biomass so if you change the endpoint in the future, it won't affect how we interpret the results, we can still use the same criteria.

RUSS MCMILLAN In the SQV report, the top 5 of those are the ones used for developing the criteria.

LAURA INOUE I don't believe Microtox was included.

TERESA MICHELSEN No it's not. Actually, the comparisons now for SQS and CSL are to the control not to the reference and I'll show an updated table.

CLAY PATMONT That's worthy of discussion. Control versus reference, and what is being used, I'm getting a little confused in terms of what the current version is.

TERESA MICHELSEN I'll show that. This is more for discussion of the task.

CHANCE ASHER: Russ could you just clarify again what the question is and what the options are.

RUSS MCMILLAN: As we're looking to develop the framework to employ in freshwater criteria for bioassay confirmatory test, we're trying to mimic what we've done with the marine standards, but we can't necessarily do that with the selection of species we have to choose from. We don't have what might be defined as true chronic endpoints and so we're struggling there whether to consider a sub lethal the same as a chronic. Easy solution to lump both sub lethal endpoint which is the 10-day *Chironomus* growth test, not the full lifecycle growth test. It's not quite a full chronic test, but to lump that with 10-day growth *Hyalella* as options for the chronic test. And then select at least 2 acute endpoints as well.

TERESA MICHELSEN Wouldn't the 28-day mortality be a chronic test too? You need to make distinction between endpoint and test, sub lethal endpoints versus chronic tests ... at least in my opinion.

RUSS MCMILLAN So you would include the 28-day mortality?

TERESA MICHELSEN But not the 10-day.

RUSS MCMILLAN Right.

TERESA MICHELSEN Yes, I might include the 10-day. We do have issues with port projects where they're on such a fast track that it's hard to do a 28-day test especially in areas like Willamette River. If you're talking about freshwater you may be talking about more of those types of environments. So I would see using something like a 10-day growth as a backup for a chronic in a short term situation

CHANCE ASHER But cleanup isn't short term.

TERESA MICHELSEN I know but these will be used for all kinds of things so.

CHANCE ASHER: But cleanup has to be our priority.

PETE ADOLPHSON I think the exposure durations are what define acute versus chronic and based upon life cycle of the organism. So I agree with you on that. I have terminology that we're going through, too, to define acute and chronic. I agree a 28-day exposure duration would be chronic for a *Hyalella*; a 10-day would be a chronic exposure for a *Chironomus* and when reduced to 10 days in each one of those, would be acute exposure and then endpoints are different.

TERESA MICHELSEN So a 10-day growth would be an acute test with a sub lethal endpoint, if that makes any sense. It's up to you what you want to put in the rule as far as whether to focus on endpoints or tests. But it's important to keep terminology straight.

RUSS MCMILLAN Yes, maybe we should just solicit these kind of comments rather than look at specific options. You caught the issues regarding life cycle in the notes. Any other questions, clarifications or recommendations regarding using that suite of bioassays we have to best to represent the sensitivity of the benthic community?

CHANCE ASHER: Or perhaps some of these questions, we could write down and they could think about it and bring it back.

RUSS MCMILLAN: That would be a good way to go.

CLAY PATMONT One comment on the 10 versus 28 day *Hyaella*, there has been a lot of dialog nationally on this. I know from EPA's perspective they might have a different view. I think they may be very comfortable with the 10 day. It's been used at a lot of Superfund sites. There are people who believe that the 10-day test just isn't long enough; it just isn't even a good acute test. I think the people making those arguments are pretty limited at this point; most of the rest of the community I hear is in support the 10 day tests, but there is a lot of discussion on these issues right now nationally.

LON KISSINGER I'll be taking some of this back to Burt Shepherd at EPA Region 10.

RUSS MCMILLAN Along those lines, the 20 days is a compromise. It doesn't stretch it out to the full 28 days.

CHANCE ASHER How about we do this; because getting the notes out will take a couple weeks. We want to turn this around. Perhaps we can capture those questions right now and then you could make some notes and have your discussion. Do we want to do that now?

CHANCE ASHER Then we'll send it out to the whole group.

RUSS MCMILLAN Homework for this group: comments on SQV report are due by May 17. In the meantime, we will be continuing to work towards drafting the rule language, setting ourselves up for the EPA and peer review and completing the Science Panel review, then it's out to formal review.

TERESA MICHELSEN OK, so as Russ mentioned this originally came out of the RSET workgroup to develop freshwater guidelines for Washington, Oregon and Idaho. I should say it originally got stalled in the work group largely because we ran out of money to do the additional studies that Ecology later funded. The SQV report in addition to all the peer review Russ mentioned is now going back to the RSET group for review partly because they're our federal partners in the dredging program; will be going to EPA, NOAA, Fish and Wildlife, along with Washington, Oregon. Idaho has pretty much told us they're just waiting for us to make a decision and will be happy to use the criteria we develop. When we get to the public review phase we'll probably send it on to them too.

So I'd thought I'd describe a little bit about this method. Don't know how many of you have seen this presentation. If you've read through it or how much detail you want me to go into or if you've read Clay's message in the attachment. The attachment to Clay's message was really good. Older paper describing all problems with existing SQGs. And I agree with almost everything in it.

One of the reasons that this method was developed was to try to address shortcomings in some of those methods. The principle thing about all those methods is that you can't improve one type of error without increasing the other type of error. It's really all up and down the scale, all the chemicals, So the brainstorm around this method was largely to let every chemical move to its' own level of toxicity, therefore hopefully minimizing both false negatives and false positives at the same time. So in order to do that, you have to work with all the chemicals at once. And therein is the trick. It's not possible to do a fully optimizing model that just runs on its own and optimizes all those variables; there are just too many variables. Dreas Nielsen tried it once and I don't know how many of you worked with him but it ran for a week on his computer before ... I had to introduce some constraints into this model to help it find its range and to help identify the areas of toxicity that each chemical would move to. That introduces some critique of the model but also helps it work. So you can look at it either way. Basically we start the same way everyone else starts. We gather our data, we QA it, we figure out how we're going to normalize it; how we're going to sum it; we do all those sorts of things. I start out with some initial range finding that has to do with policy goals for the agency. For example, if the agency tells me 5% false negatives is as low as we'll probably go, I'll start there. I tell the model- start there – find the point at which we have 5% false negatives we're going to work from that. So find some fixed percentile just like a TEL or anything else that starts at a 5% false negative. And then it allows the chemicals to float to a place where the false positives are lower and the false negatives rates are the same. So that's the one real difference between that and the other models.

The models all need their chemicals to float independently of one another but you can also run it over and over and over in different ways so you can find the synergies of the data, the chemicals that are moving together. So you can identify for example metals that are moving together that are affecting each other and can identify classes of organics that are affecting each other. You can also run this in conjunction as I've often done or other statisticians, covariance analyses or traditional statistics to see which classifications co-vary. And that helps us set up the sums and other ways to get the data into the model so they act more like independent variables. So its features are it is multivariant and it does consider all the chemicals at once. This is a pretty big deal because on something like the TELs, even the AETs, to calculate the individual value, they'll ascribe all the toxicity in the data set to that chemical, which it rarely is, it over calculates that value and that's obviously wrong. And it tends to make the number for that chemical overly conservative, particularly when you use all the chemicals together. So that's what I was trying to avoid. I wanted to do all the chemicals at once. So it does address co-variance in that way. What that does mean though is it can have multiple solutions. And this is one of the big issues we have with it. But that's just reality. The chemicals co-vary in the environment and there are multiple solutions that will work; there are multiple different criteria sets you can put out there that will work really well. People don't like that, but it's true, mathematically speaking.

But one of the first things that I do is start with a whole data set of 300 chemicals. How do we take that and make it manageable? First thing I do is screen out stuff that isn't detected; that's kind of a no-brainer. So if we don't find it in the Pacific NW, it's probably not found. We have

some chemicals that have like 30 detections out of the 900 data in the data set. So I have a list of those in the back of the report in case anybody has that chemical at their site. But the next thing I do is an ANOVA screening. It basically takes each chemical that involves a hit and does an ANOVA distribution for that chemical. So it says: what are all the stations for this chemical that got a hit; what are all the stations that had null values; what are their concentrations at those stations? If their concentrations are the same as the null stations, chances are that chemical is not causing toxicity in the data set. It's random, whether it's a hit or not. If the hit stations are higher in concentrations than the null stations, then it's a pretty good bet that chemical is contributing at least some toxicity in that data set. So that's the screening step. And I tested that by weeding out that screening step and all the chemicals that would have gotten screened out and never make it into the final model, never showed an effect in the final model. So I got pretty comfortable with that screening step over time. Even tested different "*p*" values to see which "*p*" value we should use. And the good old "*p*" = 0.05 was best.

And as I mentioned we select out a false negative target to start, we optimize the false positive against that false negative and we repeat that for the range of false negative targets because we're trying to give the agencies options. We're saying you may have said the lowest you ever go is 5% false negatives but is that what you want to regulate on? So we'll give them a range, 5, 10, 15, 20, 25, and then ultimately Ecology makes the decision about where they want to fall within the range. We evaluate the reliability every which way. We'll show you in a minute. And essentially one of the nice things is now through the RSET process this is essentially automated. There are a series of spreadsheets that more or less run themselves. Mike Anderson at ODEQ did this for us for which I'm eternally grateful. So I no longer have to run this model every time it needs to be run.

This is just a little diagram of what I mentioned. The blue line is like a 5% false negative line which is where everything starts. If you look at chemical "B," the toxic range is the red range and so chemical "B" really can't move up at all because it's already toxic in that 5% false negatives. That's why there are some false negatives at that level, some metals for example are already toxic at that level but they get pegged there and have to stay there. Others like chemical "A," the toxic range starts well above that level and so it's allowed to rise until it hits the red and then it stopped. So every chemical is allowed to rise to the bottom of the red. Because if it went above the red then the false negatives will start to increase and we don't allow that. So each one just floats up to the level until it hits the red. Now the fact that there is overlap between the green and the red means there's an area where sometimes it's toxic and sometimes it isn't. It depends on the bioavailability at a given station. Some chemicals have more overlap than others. Some have a pretty distinct boundary, at least in this data set, and that's data set-specific. Others have a big wide range for that, it's true. So I just depicted some of those differences here. Look at chemical "D." Some chemicals will move all the way to the top of their measured concentration and never show toxicity. You have to say to yourself well maybe those chemicals aren't really toxic. Those are the ones we'll see later, with a greater than value which originated with the AETs, and basically says the actual toxicity level is unknown but it's greater than this level.

So this is a text description of the different reliability measures we use. Look at page 14 of the report if you have it with you, it shows a visual diagram which is much easier to follow. It shows the relationship between actual toxicity and predicted toxicity. So a false negative is when you say that something is not toxic, but it actually is. A false positive is when you say

something is toxic, and it's not. And overall reliability is how many of those you get right. Whether you call them toxic or not, how many do you get right. The other two are a little harder to describe. It's much easier with the diagram. But it's basically, as opposed to what's really toxic or not, it's of the predictions you make, how many you predicted are right or not. But that's what the regulated public sees. Nobody knows what's actually toxic or not once you're done with the exercise. We know because we have the chemistry and biology for each station. But when you go out and start using the SQGs, you aren't going to know whether those stations are really toxic or not if you're not running bioassays. So the predicted hit and no-hit reliability has a lot to do with the regulated community's perception of whether our SQGs are accurate. We are using them to predict a hit, or predict a no-hit and then they go to run bioassays and they see how it matches everything. So we use all these measures for reliability for both the lower and the higher levels in fact which is a very big difference from the national paradigm that says that the lower level only has to be able screen stuff out and the upper level only has to screen stuff in and in between you always have to run a bioassay. Well guess what, that in-between is huge. It's most of what we deal with. So we tried to do something different here. We tried to make both our lower and our upper level have good reliability. We think that's very important.

So our overall data set increased substantially from 2003, it's a pretty robust data set. We wish we had more chronic *Chironomus* data, that's why I recommended we try and get more. We looked at a lot of issues over the last 4 years. We looked at geographic diversity, we looked at different measures of petroleum hydrocarbons, whether or not to include Microtox, what to do about the crazy *Hyalella* growth data from Portland Harbor, which was eventually excluded. Some issues with ammonia and sulfides data from Portland Harbor, some weird stuff in the data set. And qualified pesticides which came from Portland Harbor. We had to decide whether to include it or exclude it as it affected like 200 data points. Had to re-blank correct all the data because we discovered 80% of the way through the surveys were not blank corrected in the same way and it really affected the data. A lot of the lab contaminants were showing through in our model and they had no business being in there. And when we properly blank corrected them, they magically disappeared. We were very happy.

We looked at the control versus reference issue several times. We did it in 2003. Did it again when we had more data because we thought well we've got more data and once again there's no benefit to going to reference; there's no added reliability, and you lose a lot of data. So we stayed with it. But it did lead to revision of the bioassay interpreting criteria indirectly, the fact that we were using control vs. reference along with a lot of other factors. So revision of the bioassay interpretive criteria is the most recent thing we did with this group and with national experts in the labs. So let's talk about that.

This is the new revised table and originally both of the 10-day mortality tests had 10% difference from the control as the interpretive guideline for SQS; that was really stringent. And even the recent workgroup, well perhaps that was the problem. These two tests were coming up as being somewhat unreliable. All of our other tests met the reliability goals and we accepted them. These two did not; that was really the big hurdle in getting these SQVs approved. So when we came to this work group we heard exactly that same comment, that those levels were probably too stringent. We went to our national experts and heard the same thing; we went to the labs and heard the same thing. So we said well let's test it statistically, so we did that. We tested a whole range of different differences from the control. And we picked the ones for which the ANOVA test showed the strongest statistical correlation between the chemistry and the bioassays. And

also the ones under which the reliability improved the most. So those were our two criteria. We tried to make them as objective as possible. But as you can see they're still low within the range of all the other bioassays. It's not like they're suddenly outside the interpretive guideline range for either the fresh water tests or the marine tests. And also you'll note we're now comparing to control rather than reference which is a more stringent comparison. So we felt that loosening up on this comparison a little bit was justified. In addition we may be tightening up these control guidelines. There's a move nationally to tighten them up a little bit because labs are actually getting better performance than this. So if those get tightened up, then the overall mortality in the samples; tightening the control and loosening the comparison overall mortality will be the same end result. And I think that works pretty well. By the way some of the other endpoints are also getting revised. There's a debate going on, and you can let us know how you feel about it, over whether to revise these main criteria now or wait for ASTM. I think it's kind of good to let it go through the whole process.

Just to show you a little bit of a typical reliability result that we got. Blue shows the line that met our work group goals for reliability; light blue within 5%; yellow is designated below. And then below we compared it to the other existing guidelines that are available. So this is for the SQS level, the lower of the two levels. And for every bioassay now that we've revised the interpretative guidelines, we do have a dark blue line. We have a line that meets our work group goals, which was below 20% for false negatives and false positives and above 80% for overall reliability, that's what we wanted to see. If you kind of look down below, you can see, if you want to be fair, all those false negatives are in the 5 to 10% range. So if you look at the 5 to 10% lines up here and compare those with the ones down below you can see that they have almost double the false positives of the similar lines for the FPM. The overall reliability is only about 1/3 of the FPM so we feel like we're making a really significant improvement on the other alternatives that are available regardless of how conservative you want it to be as to which of these false negative lines you chose. So Laura took all the different bioassays and averaged them and I stole this slide from her and this just gives you kind of a general sense of the floating percentile methods vs. the other approaches that are available and how they compare. So the differences are really huge at the SL1 level which is our overall long term goal for Puget Sound. Also governs the dredging program so it's pretty important for that reason. And then for SL2, you can see that all three measures are improved over the other alternatives that are available. Laura is going to show it again at SMARM in a couple days so if you want these slides they'll be available. I guess we'll post them.

CHANCE ASHER Yes, these will be on the web site under the work group.

TERESA MICHELSEN So how did we finally go about selecting the criteria? And here's where I can mention a little bit of a conundrum that the work group had which we've since resolved. And that is, this is what we expected. We expected that we'd derive a bunch of SL1 which is the lower value like an SQS and they'd all be sort of clustered together and then we'd derive a bunch of SL2 and they'd all be clustered together. And it didn't work that way at all. Basically each bioassay was more different from the others than the two effects levels were from each other. Now if you've ever worked with LOEL data, you've probably seen that before with LOEL data so I guess this wasn't too surprising when we saw it. But it did result in a real conundrum in how to set the final numbers. So we finally decided basically that the rule says the values should be in the range of no effects to minor adverse effects and that we would create a distribution of values from no adverse effects to minor adverse effects so that's what we did.

And this is the distribution for each chemical; it's in the report. The blue, greater than values, are what we talked about before where for some of the bioassays the toxic levels above any of the levels in the data set and over on the right is the proposed values. So the SL1 SQS is always the lowest value and the SL2 CSL is the next highest and a different value, which requires a tad of explanation. Sometimes the next highest value was almost the same so the red on the left here shows a really good example of that - 0.54 and 0.64 aren't real different from an analytical perspective. We wanted to leave site managers some range within which to work so instead of choosing 0.64 as the CSL we chose 1.7 as the CSL. So whenever the differences were really small between the lowest two values we went up to the next highest value. It's still pretty conservative given that the entire range is no adverse effects to minor adverse effects and theoretically we could have been anywhere within that range but chose pretty much the lower two values. That's partly because the model itself isn't usually conservative, it tries to be realistic. We're being conservative at this end and we're also responding to comments from ODEQ and EPA and other agencies that are more conservative in some ways so we're trying to make sure we have consistency among agencies and all of our different regulatory authorities apply all at once.

So next steps, we'll continue peer reviewing this, we'll have the Science Panel report, we'll complete the draft rule language, we'll have EPA review and formal public review and RSET already has this or will have it soon. When we sent it to ODEQ we gave them the go ahead to, because Keith was the chair of our RSET work group, we gave him the go ahead to distribute it to that work group at the same time that you all got it. I don't know if he's done that but he will soon if he hasn't yet. And that's it.

CHANCE ASHER OK

TERESA MICHELSEN It's been awhile since I've had to present that to a bunch of technically smart people so feel free to pepper me with difficult questions.

CHANCE ASHER Yes, so we should just open it up. I'm not sure Teresa if you had some other specific questions that you want to pose or if we just want to open it up to general discussion.

TERESA MICHELSEN I just basically came prepared to answer any questions that I could or if you have comments to give to Ecology. But I'm here basically as a resource for you to understand what we did and then what you think.

CHANCE ASHER So any general comments or clarification questions?

CLAY PATMONT I think the biggest question I've got is, and it's probably more for Ecology. Is just how reference comparison, I understand what you've done with the control from a derivation and chemical screening criteria. I think you've actually you've done a very good job but there are some details. But the biggest question I have is just for a confirmatory test, are you advocating a comparison to a control? So you're in this zone where you decided you really need a site specific bioassay or a level in the gray zone, it's not clear when I read the report or maybe I just haven't looked at it closely what the recommendation is.

TERESA MICHELSEN - The report doesn't make a recommendation, that's why it isn't clear. For example the Corps of Engineers has always been comparing to control for a long time -so it will just do that. What Ecology chooses to do, I don't know. I know that we did try - both RSET and Ecology attempted to find good reference locations. We went through about a three-year process of doing that. We largely failed. I hate to say it, but we failed. That doesn't mean that some sites haven't been successful doing that. So this seems like a really good site-by-site flexibility kind of thing to me. I also think that for Ecology's reference area report we should plan to go back through with the new bioassay interpretive guidelines because I have a feeling some of those phantom bioassay hits may disappear with the new guidelines because they were just too stringent. So if that's the case we may have some good fresh water reference areas, we just haven't done that evaluation yet.

RUSS MCMILLAN I think one of the challenges is as you look at applying the standards like these in a geographically wide area, you're going to just struggle to come up with appropriate references. If you've got a site like Portland Harbor where you have site upon site adjacent to one another and you've used the area extensively, or sampled it extensively, and come up with what might be reasonable reference locations, it perhaps represents a different picture than applying the same choice across the state. So we would welcome any comments you might have.

CLAY PATMONT We've talked about this so much Russ, in the context of the marine larval test. But you're setting forth QC requirements for reference areas and yet I think we all know the reference pool is highly uncertain, so you're basically setting yourself up for a situation of rejecting. You know I think there's ways to clarify but it just seems like you're inching towards a pretty dangerous zone of setting reference criteria that by all accounts, you don't really know whether you can actually meet them or not.

RUSS MCMILLAN Yes, right.

CLAY PATMONT So that would speak to false positives. That ends up, you have to reject your reference pools you end up having to default control comparison. And yet, there may not be anything going on with the site but you have a hit. Now you've got a situation where you have to explain some toxicity and it may be totally an artifact of the test or the conditions but may have nothing to do with the hazardous substance that you're focusing on.

RUSS MCMILLAN I hear what you're saying. I guess one of the challenges there is, well I think part of what you're interpreting as direct indications of, reference QA/QC requirements, is perhaps just a carrying over use of tables and what not that have been derived from the ASTM work to begin with. And doesn't necessarily represent our expectation that we will have those for reference. Which means if we choose to default to a control, then your comment holds. Are we looking at true toxicity? It's always a challenge.

TERESA MICHELSEN It also means they don't put those reference criteria in the rule but you might have some flexibility if at a given site you decided to do a comparison to a reference, or to have flexibility at a site to set up reference. If it's this new reference envelope method, these don't even apply. You know it's a completely different statistical method. So a lot of the reference discussions, as I understand it, nationwide are moving toward that type of approach and away from the sort of traditional T-test type comparison. So it may not be smart for us to

codify that sort of direct material approach to the rule. But, you know that's a question we've never talked about it. I don't know whether those would end up in the rule or not. Or in the SAPA guidance or similar type of guidance.

CHANCE ASHER Well we haven't really discussed that. I don't know if Ecology people have an opinion. We want your comments.

CLAY PATMONT I see it as a huge issue obviously. It's and, I think what you've done is really good on chemical criteria, I mean it makes a lot of sense and the thing I worry about is just co-opting all those tables and saying oh and here you go, here's your reference. Not all site managers are going to have and are close to the level of understanding you have of this site. You're going to have people making calls.

TERESA MICHELSEN I actually see this as one of the big advantages of promulgating these values because last time around we didn't, and the report itself became a sort of something that site managers would go to and look at. If these criteria actually are promulgated and then the pieces will be clear what Ecology intended to promulgate and the framework that goes around it and then the report will just kind of become an historical document. You know what I mean? Like the old days, these reports they won't be relied on so real time by site managers and consultants and be less chance for abuse of tables that aren't intended to be the law of the land. Questions? Anything I talked about that didn't make sense, or you didn't understand?

CHANCE ASHER It might be good for us to bring up some of the comments that we've heard either from internally or from ODEQ so you guys understand what the responses are because you may be asked some of these questions as well. There are some questions that Teresa had mentioned, and I need Russ and Teresa and Pete and Laura to speak up too. Teresa had mentioned the fact that the false negative rate doesn't change, it will stay at 5% and then you can run your model. But there are some issues that folks have with adjusting the rates. They have issues with the fact that you can do that and that we're not necessarily choosing the absolute most conservative rates. And I think that would be good to kind of get that out on the table and talk about. Because it's a policy issue more than a science issue that we're looking at making but it would be good to get feedback from folks as well.

And I'll kind of start it off. One of the things that we're trying to do with these numbers is number 1, try to stay similar to the framework that we currently have in the SMS rule and not use an ultra conservative framework that we haven't used for the marine criteria.

And the fact is that we're looking to promulgate biological criteria as well which would be your confirmatory value, or confirmatory test, if a site manager suspected that the number that we chose for promulgation wasn't protective enough. We didn't feel that it was necessarily the wisest choice to screen everything in if you have such low rates. So does anyone else want to speak up? Russ, Pete, Teresa? But Russ and Pete for Ecology policy issues, and then Teresa for some of the technical issues with the floating percentile method.

RUSS MCMILLAN - It has been quite a challenge trying to address the idea that we should be protecting and winnowing out absolutely where there is no toxicity as opposed to setting that lower range at some acceptable level of false negatives. And our operating assumption is that we would start at that 20% for what we would look at as a choice and then balance that false positive

rate to as close to that 20%, below that 20% as possible. When you're comparing these numbers to consensus-based numbers, there are a lot of numbers that are a lot lower than these and would sweep a lot of sites or stations in, identify them as toxic but when we go to our confirmatory biological test there's a good chance that those won't be toxic. So it would tend to cost a lot of additional resources and money for both the regulated community and the regulators. So we're trying to find that best balance of false positives and false negatives.

TERESA MICHELSEN There's a couple of technical and statistical population notes on this too, that can be brought to bear. And one is that our data set, which I believe to be representative of the types of sediment data we'll be collecting in the future because it covers 15 years of program history now, has about 15% toxic stations and 85% nontoxic stations according to the bioassays we've run in the past. If you start adding a bunch of chronic bioassay tests in the future that number could go up a little bit, but let's say 20%. What that means is that a false positive is a much larger number of stations than a false negative on a statistical standpoint. So if we have a large percentage of false positives, we're sweeping hundreds and hundreds of stations. Where if we have an increase in false negatives of 1% or 5%, that's a few stations. So there's a big difference there. We never make decisions on sites based on a single station. So keep in mind that this is a per-station error rate and we make decisions on sites based on a minimum of three stations according to the rule. So, the chances of making that error at all three stations to designate something a site when it shouldn't be, or not designate a site when it should be, is the multiplicative error. It's  $20\% \times 20\% \times 20\%$  which is well below a 5% chance that you will fail to designate a site that should be designated. It's a small chance. People will never think about that part of the framework that this is going into. So it's important to realize that there are other fail safes on top of a per-station error rate. There's a whole framework that it's in.

For a dredging project it's a little different because you might be making a DMMU by DMMU decision but you're also layering other substance on top of it. That's up to them to worry about.

PETE ADOLPHSON Yes, we've talked about that as well and I had thought about that more from the last time when you brought that up and that's a really important point. What I always hang my hat on is, and that's from site management point view, is that there are certain sites particularly in freshwater that have unique situations. And you can look at the buffering capacity of the water body, you can look at whether or not you have extreme seasonality in that water body. They're unique to the water body itself or at least the location of the water body. And if that's the case we're still following the same paradigms that we do in the marine. And that's where the importance of the bioassays comes in. If you have reason to believe that those numbers are not applicable, Ecology or the PLP can run bioassays whether the chemistry shows it's a hit, not a hit, or the site manager suspects the numbers are not protective due to site specific conditions. If we don't see contaminants that exceed our values but we have reason to believe they're bioavailable at lower levels, we can require bioassays. And that's why it's important to have a balance so that you can weigh that; it's not a mandate for cleanup one way or the other, but you balance whether or not you need to do it based on site-specific information.

TERESA MICHELSEN I think this is critical for freshwater guidelines more so than for marine because we do expect these sites to be more different from each other than our marine sites, so the more bold we can be in our rule about being really clear about Ecology has the right and the PLP has the right to use biological testing approaches. I think it will be really important to do

that. Also this is when I get a little concerned about trying to run validation tests. This is an issue that's come up over and over. Well we got to validate these before we promulgate them and we don't have a data set with which to do that yet. We will hopefully eventually and of course we'll do it. But we're going to have to keep in mind the idea that there's going to be much greater variability in these sites. Well what we can't do is pick one individual site out, especially if it's a mining site or some other site where there's weird chemistry. To say it doesn't work at that site, it's all screwed up; we can't use this. And that worries me. I have that worry about some of the comments being made right now.

LAURA INOUYE I remember when we were working gathering the data we got really frustrated with at least one batch of data. And I'm wondering how many there are that we couldn't get the QA2 level of validation. Was that more than one project? Were there a couple? Because I was thinking maybe that could be used as validation.

TERESA MICHELSEN - There were. They were pretty small. There were only a few stations in each one which is why I didn't get too wrapped around the axle. But yes there were some labs that just basically said we were never asked to create that in the first place. It would be a huge amount of work for us to do it. We'll do it if you pay us. And there wasn't any money or time to do that. So for some projects we just weren't able to get that.

LAURA INOUYE - And I was wondering if there were sufficient of those and the data that's going to be coming in to start to use as a gathering of paired data for.

TERESA MICHELSEN - Yes I think it's worth gathering it. I still would want to see it be regional and a substantially sized data set before we try it just because of this variability issue. If we pick out an individual site and it has different chemistry or whatever. I mean the goal here is to be reliable on a programmatic basis and then if you have specific sites that are different as Pete was pointing out we've just got to be more aware of that and more flexible than with our marine sites where we really expect them all to behave the same way.

RUSS MCMILLAN One of the things that we're looking at including in the language would be some precautions to take in terms of assessing whether your site is likely to be representative of that same data set that these criteria were developed from. And that really gets to the mining sites and some of these that are peculiar because of the bioavailability issues. And it will be a challenge to write something that goes beyond. Everybody has the right to confirm with bioassays. Because we don't have ranges of pH or hardness or some of the other things that might come into play there. But we do want to capture that in very explicit language that this may not be universally applicable across the state.

TERESA MICHELSEN And you can have more guidance in the SAPA. But you know, I don't know if you like the current ODS section has examples in it and that was pretty bold I think to put those examples in there. But they've been useful over the years to say hey this is what Ecology meant, as your deleterious substance. I think you should just go ahead and put in a couple examples and you can't give pH ranges, or acidity ranges or anything but you can mention the pH or acidity and unusual matrices such as may be present in mining sites as examples of cases.

GLEN ST AMANT I still have a question. I apologize if this isn't the right time to discuss it. I didn't have as much of a question on the SQV derivation but kind of taking a step back and looking at those numbers and applying it to fresh water sites. How do you think about other eco receptors at the other sites? Is that totally going to be site by site?

CHANCE ASHER These numbers are definitely set the same way they are for the marine criteria which are to protect the benthic community from acute and chronic toxicity and so it's not meant to cover the bioaccumulative impacts and other receptors. So that will have to be folded in with the bioaccumulation issue that we had discussed on January 28 and will discuss again on June 2. But it's still a feasible way to structure the rule because we currently have that in the rule for the marine criteria but we just need to have some clear language about the bioaccumulative piece. So it doesn't really impact these numbers, and promulgating these numbers, it would be a narrative, but you'd have to look at all exposure pathways and the bioaccumulative impacts as well. Laura?

LAURA INOUE Working with the RSET group (Regional Sediment Evaluation Team), it's been an acknowledgement that this is benthic organisms only and that fish and other nonbenthic and invertebrate creatures may be not just bioaccumulatively, but toxicity impacted, we were working with NOAA Fisheries to try to get some values on that. There was some progress made but not sufficient enough to say here's some numbers that are protective of juvenile salmonids. There moving in that direction, but they have a lot of time constraints.

TERESA MICHELSEN Whatever happened to that? I thought we were close?

LAURA INOUE They found the area where you started seeing the bend in that model where you start to see adverse impacts. And it became a policy issue of is it really where you start to see a deflection from no effects into the model shows you start to see increase in tumors and other various endpoints in English sole. And there was a matter of, well how much, is it right at the inflection or is there some error bars associated or certain percentage? And it flat lined from there. The scientists said here is the model and here are the values and we were unable to get more of an overall policy from NOAA from where they felt comfortable saying this is where we think you're going to start seeing impacts to the salmon or the other fish. And that was the PAHs. I know there's been some work done with DDT and those compounds as well. It's much further behind. There was some stuff going on with concentrations and stomach contents of salmonids that were associated with toxicity seen in those juveniles. That was even further behind. So it's just, it's not quite far along enough to develop.

GLEN ST AMANT So it's just going to kind of be a narrative place holder?

TERESA MICHELSEN We asked them for those two levels, you know the sort of SQS equivalent.

GLEN ST AMANT And don't misunderstand, I'm not asking you to develop things that were previous.

TERESA MICHELSEN I'm just frustrated.

GLEN ST AMANT I'm just trying to figure out what it's going to look like.

CHANCE ASHER Yes, well at this point based on the conversation that we all had on January 28, we're still working out what that language will look like and we hope to have that ready for you so we can have more of that discussion on June 2. So we can tie both this and the criteria issue together. There are a few other things I want to point out about these numbers versus other numbers that are out there. There are other criteria that we could have adopted if we'd wanted. But the one thing that we wanted to do is stay as consistent as we could to the current framework we've got in the SMS rule. And then also we like the idea of the numbers being developed from regional data rather than looking at nationwide data or data that hadn't necessarily been QA/QC'd appropriately or had been taken just directly from a laboratory. This is data from the real world and we thought that was really important, and it's regional as well. We've had some concern that there may be a difference east of the Cascades versus west of the Cascades for fresh water. At this point we don't really have a good handle on whether that's truly the case. I think that like Russ had mentioned, it may just simply be the types of contaminants in that environment rather than the fact that it's one side or another of the mountains. And we referred specifically to mining sites that are predominantly metals. So we weren't looking to necessarily separate east and west.

TERESA MICHELSEN - I actually did test it. I did runs that were East and West separate and then combined and the combined had better reliability than either of the separate alone. So I don't know whether that's an effect of having a larger data set or why it's better combined. Maybe if we had more equivalent amount of data from the east side we could better look at that, but so far at least there's no statistical reason to separate the data.

CLAY PATMONT The one, I guess this is more of a technical question for you Teresa, the selection of using dry weight balance, the consistency with the marine program. A lot of polar organics. I'm interested in your thoughts as to why you changed the program, just have everything on a dry weight basis. And then there are specifics like a TPH and you use like a diesel range. And is that really a parameter that we feel is representative.

LAURA INOUE Actually, I think very much so. There are a lot of PAHs that are not in the 17.

CLAY PATMONT What's the best measure of that oil mode of action. And I'm thinking more VPH/TPH and those other tools.

TERESA MICHELSEN Right. OK so those are two totally separate questions. Let's take one at a time.

CLAY PATMONT So it's more of a selection of the hydrocarbon, so what's the basis for that? The metals or some other AVS SEM is a more refined methods. These are concepts that have a lot of traction, a lot of history. Are you thinking of going down that road and then that opens up maybe like a half way step between the bulk of criteria and doing assessment of bioassay and looking at pore water. And again huge advances in the technology in the last few years. If we were to have this conversation 5 years ago we wouldn't be able to say what we can now measure. The rest of the country is going down these roads.

TERESA MICHELSEN No, not all of them.

CLAY PATMONT - I know. But there's a lot of good information. So I guess I'm kind of tossing out all these in relation to bioavailability. Have we selected the right parameters to assess the relationship of the chemical?

TERESA MICHELSEN Let me give you a bit of history and then Ecology can add to what their opinions might be on this. RSET looked at all these things not only for our work group but the chemistry. So the first topic is organic carbon normalization that you brought up. We have looked at that 5 or 6 times. We looked at it for Los Angeles Harbor and San Francisco Bay, we've looked at it for the marine AETs.

So back when we did the marine AETs, we did them both ways and there actually was no difference in reliability at the time. The only reason they were organic carbon normalized was largely that they were the theoretical preference on the part of those who were deriving it. There is no real difference with the reliability with the two sets of values which is one reason PSDDA uses the dry weight values. You know the rule has the organic carbon normalized values, but either way they can be used interchangeably. During the rule implementation there have been issues with organic carbon normalization over the years. Some sites have too little organic carbon to fall within the upper limits of the partitioning zone, others have too much. Sometimes it is anthropogenic in origin. It's in a weird matrix and nobody knows what to do with that. Sometimes people just don't understand it. I had to be an expert witness at a trial where the US Navy attorney didn't even know what it was and totally misrepresented it. There's a lot of just public understanding issues that come up with those organic carbon normalized values. They are just hard to describe. We don't have any problems with them. So given that there were implementation issues and they're not any more reliable than dry weight values, so then the question is so how many other times have we looked at that? Like I said, we looked at it for LA, for San Francisco and looked at it for the fresh water guidelines twice, both in 2003 and again in 2008. Neither time did it show any appreciable difference. And there's no other set of SQGs nationwide or any other country that I'm aware of that's actually organic carbon normalized. Not including partitioning values which have actually never been promulgated. So we're pretty comfortable that sticking with dry weight is a good approach. But that's just the history of it. We couldn't see any value to it either from an implementation standpoint or an actual liability standpoint. Theoretically we all know it should work but apparently it doesn't.

LAURA INOUE At SETAC last year there were several talks about the organic carbon and what's really going on. Because scientifically when you think about it, organic carbon normalized should work better. But what they found is there are at least five different types of carbon that are all having different partitioning. But I think it just gets so complicated and because at each site you might have a wide difference in the type of TOC much less, is there enough or is there too much? And there's so much noise in it when you try to organic carbon normalize and run those models, it just falls apart.

CLAY PATMONT That's all good. Sounds like you're saying it doesn't make a difference whether you carbon normalize or not so you just pointed to the dry weight. But yes the same five different kinds of carbon. Things like the soot carbon, black carbon being so darned important leads to towards a sort of pore water base measures. Right? That's where, again that's where a lot of folks are going with this. Maybe it's more of a policy issue. These are useful for first screening. What's the next step? Have you thought about the next thing?

LAURA INOUE The pore water I think might have some very good relevance with the metals. I think metals could be greatly improved if we looked at pore water. The organics I think you're going to start running into problems of detection limits and also getting enough poor water to run all the analyses to get the good low detections limits that you need. And there's also some other questions on that on the real availability of those compounds. It is simply pore water or is there some other things going on with the organisms ingesting sediments. I think more so than for metals the various surfactants and other protein degrading enzymes in the stomach contents of these critters that are eating some of the particles. There's been a lot of research showing that really helps assists the desorption of organic contaminants from the organic matter in the soils and sediments.

CLAY PATMONT You are trying to predict *Hyaella* and survival and growth and the same thing for the *Chironomus*. Trying to come up with a better measure for that endpoint, because that's what you're using for your test. There's all kind of debates on all this stuff. I'm just kind of thinking practically. I think this is really useful as an initial screening and if your site's lucky enough to be below all that stuff, you're saying great. But what's the next step? That's where I'm sort of getting to.

CHANCE ASHER - Russ or Pete do you guys have a response to the pore water issue or the TOC?

RUSS MCMILLAN If you've got good indication of where SEM and acid volatile solids or sulfides really play a good role in developing criteria. Bring the stuff to us.

TERESA MICHELSEN That's been brought up throughout the history of the sediment program as an issue. First of all we didn't have any data to use to calculate it with, so we couldn't have used it to calculate our guidelines. But the reason that's the case is that historically site managers as you've probably encountered haven't been that receptive to AVS SEM. Part of the reason is it only works for certain metals, it doesn't work if the area is disturbed, it has to stay anaerobic. Well, we've got all kinds of site uses that disturb and oxygenate sediments and there are all kinds of reasons why AVS SEM again doesn't always work in practice as it's theorized in a laboratory or in theory to work. I think that in the real natural environment they are constantly being stirred up and moved around, anchors dragged, ferries, etc. There's stuff happening that destabilizes those compounds. So we haven't tended to, as a group, use that much at sites in the Pacific NW. I can't say that I've seen it used a lot elsewhere. I think if you have a site with specific metals that were really targeted by that theory and where it really was anaerobic and was a site-specific condition that really met that, it could be a great RI/FS tool, and site-specific flexibility.

PETE ADOLPHSON And just to add to that, I don't know if you're specifically talking about it in relation to fresh water systems only, but I will say that, we did, when I was working in Chesapeake Bay, for 10 years we did AVS SEM on every sample that we ran bioassay on and it bounced all over the place. The theory sounds good which is why we did it, but we never saw any relative trend that we could identify.

CLAY PATMONT Right, I remember talking to some of EPA's folks who have looked at this a lot. It is mostly a mining site, maybe lakes that used to be hit with copper sulfate really hard, you'll find them. I talked to Dave Mount about it a couple years ago. He said he's never seen a

freshwater sediment site with metals toxicity. You know, and I think that's part of the problem. Because it is locked up and it's not the sulfite, sometimes it's the carbon, there's a number of ways that you can sequester, but it's these mining sites I think those are the ones where it pops up. And I think the Spokane River seems to be a pretty good example of ones where we do have over the amount. I don't know that for a fact. These are tools I guess. I don't know if we have to address it here.

TERESA MICHELSEN I don't know if this reflects your comment, but it's some of the comments we're getting about these freshwater guidelines are that people are used to working with TELs and so when they see the values that the model gives for metals, they are higher than they anticipated. And that may be a reflection of what you just said but it's not clear that we can predict it any better than we're doing with our bulk sediment measurements. But that is one of the issues we're encountering is: people are expecting the metals values to be lower than they are. And just because they've grown up with the TELs and so those are the values that are fixed in their brain. Going to an important issue is petroleum. This is one of those topics we ran every which way. We did individual PAHs, we did summed PAHs, we did organic carbon normalized PAHs, we did molecular weight normalized PAHs and more fractions, what I mean to say, more fractions, we did TPH. We did not have EPH/VPH to use, so that wasn't an option. Of all of those methods that we've looked at for the last 8 years, TPH did by far (double) the reliability of the other methods. And I think it's due to the main factors that it include aliphatics as well as aromatics. I think that aliphatics are having a more toxic effect on benthic organisms in particular. It's a much higher level effect on benthic organisms in particular.

LAURA INOUE There's the whole alkylated aromatics as well.

TERESA MICHELSEN And there's the alkylated aromatics which are included in the other PAH measures and we don't have any other way of measuring them otherwise. So when we use bulk petroleum measures, we use all that. And it had better reliability despite the fact that only about half the database had data for it. Which tells you that it's really a key measurement to have in there. We did look at just using TPH and not even having a summed PAH measure in there. But the problem was we did only have TPH for half the stations and so we lost a little bit of reliability but it was tiny compared to the TPH contribution. In the end we had the best reliability including both of those measures, the summed PAH and the two TPH measures. We could have added gasoline but we didn't have enough detections. It got screened out. So what we have is what's called diesel and what's called residual but we all know that it's not exactly diesel. And then the chemistry group looked at for awhile, we decided that even though the historic data, and this is despite the fact that the historic TPH data wasn't really in cleaned up columns or anything necessarily. It had all the warts and problems of all the old TPH measurements and it still worked far and away better than the other approaches. So there is a recommendation to clean up the methods. So in the future we're really just getting site contaminants and not decomposing leaves and whatever else might be out there. But when we met as a group of agencies in the summer of 2008 to talk about this exact question, what in the future we should use ideally? The VPH/EPH did rise to the top of that discussion as gosh we'd like to have that. But Todd, I'm going to ask you to take over this discussion. (RSET continued to look at that.)

TODD THORNBURG (Anchor QEA and RSET chair of chemistry group) As a follow on to the sediment quality discussions, one of the things we did was compile all the information we could

find where we had sediments. We've been involved in the sediment guideline discussions because of the importance of getting valid chemistry to support the database. But one thing that we did was compile all the information we could where sites had analyzed both TPH and EPH/VPH where they had both information and just did some simple correlations. There's no EPH data, virtually no EPH data in sediment sites that Teresa can use in her existing database. One thing we wanted to look at was whether we could look at developing a guideline for TPH and make some extrapolation and see if we might be able to predict what a correlated EPH guideline might be. And just quite frankly, the correlation was lousy. A very low correlation coefficient. There were orders of magnitude errors both ways and so it had very little predictive capability. I mean it's good in theory. EPH gives you more information given that narcosis is based on mole fractions more than just straight up concentrations. In theory it would be potentially a more valuable analysis, it is also a more expensive analysis and there are fewer labs that can do it.

TERESA MICHELSEN Aren't there some analytical issues that ultimately resulted in a recommendation against it. That's what I was trying to recall.

PETE ADOLPHSON We had detection limit issues with EPH/VPH vs. NWTPH at the Skykomish. But you can't meet the same detection limits.

TERESA MICHELSEN Did you guys look at that? I can't remember.

TODD THORNBURG That's a little fuzzy in my mind too.

TERESA MICHELSEN It's been awhile hasn't it?

TODD THORNBURG Yes, it has been awhile. There are situations where you run EPH and VPH on a sample and you get the same bin in both fractions and the bins don't match up. You know, the lighter EPH and the heavier VPH by the two different methods don't always give you a consistent result. So that's a little trouble. But TPH seems like it adds a lot of reliability to the database and we just don't have any additional information to move to EPH at this point in time. I think the thought was very large cleanup sites that you know have expensive cleanups on the line may find some use in doing some EPH analysis to better understand their toxicity but as a global recommendation, we're not there.

TERESA MICHELSEN So it's weird. It's another one of these cases where we think something should be better one way, but we keep finding that more of these actual on-the-ground reliability comparisons we do, that simple approaches work just as well as theoretically complex approaches. I think it's largely due to the noise in the environment, the real noise in the environment. The environment is just not as well behaved as our models predict it should be and so our more bulk measurements, whether it be just plain old dry weight sediment chemistry or TPH or what have you, work just as well in predicting toxicity as the more complex theoretical models do. In some ways this makes it easier to implement. At least we've gone to the trouble of demonstrating that. That's where we're at with it.

ROB PASTOROK (Integral Consulting) Well maybe just to follow up on that, and I think this was said before. But the noise may be due to cross section of data covering lots of different heterogeneous sites and contaminants and the option of having the more complex approaches or

at a site-specific basis, could be valuable in the future in developing understanding of not just those approaches but also how the simpler approaches are behaving across heterogeneous sites as well. So putting some language in the rule along those lines to allow people and maybe even encourage them to do site specific approaches if it's appropriate. There may be places where SEM EVS makes a lot of sense or TOC normalization of site-specific data makes sense. And the other thing in discussion of the bioassays, but I think a lot of times we try to make these tools a lot more precise than they really are and that's why we have several of them. But what I didn't hear and I'd like to encourage, is just more comparisons with field evaluations not necessarily into full validation exercises. But how are these current bioassays tracking with or correlating with the species out in the field that are tested in the bioassays or other species, that bioassays species are supposed to be surrogates for, and how are they correlating or tracking with community level endpoints in the benthic community. That's important.

TERESA MICHELSEN We did go looking for benthic community data. We found zero for freshwater. We did though find a bunch of Index of Biological Integrity (IBI) data, but that's not very helpful and even most of the IBI didn't have chemistry with it so was even less helpful. I don't know. I think people have largely given up on benthic community analyses in the heterogeneous freshwater environments that we have. You know if we were the Great Lakes they do a lot of it there because it's more consistent. So I don't know.

PETE ADOLPHSON Yes, I've had that same experience in the past. I mean those approaches in freshwater have been put forward by PLPs and it's not out of the realm of possibility to do it. But because of the heterogeneity, seasonality (between seasons) my approach to them in response was I think we can do that but it probably would take approximately 40 samples both reference and at site every season for about 5 seasons. That's the heterogeneity we're dealing with, and obviously cost effective and it's not.

TERESA MICHELSEN Places like you know Lake Washington or Lake Roosevelt that are big enough and consistent enough that maybe could be done.

ROB PASTOROK I'm not advocating using at every site, I'm saying that in select cases there may even be special studies to help bolster confidence in the bioassays or to at least know how they relate to the field conditions.

TERESA MICHELSEN Yes, we sure got that comment and I wish we could have run some data to look at it.

TODD THORNBURG Really good work Teresa. I applaud your diligence. It's been many years I think.

TERESA MICHELSEN 13 years in fact.

TODD THORNBURG And you know I think this is a big step forward from the existing SQV paradigms that are out there nationally. I'd like to focus on one particular chemical. I hope I'm not being too detailed but I do think it's going to be one of our big challenges as we're looking in the decades ahead. These long term recovery goals, bay wide long term recovery goals. And that's bis-phthalate. I don't understand how the numbers are working out and I don't know if you're able to provide any insight based on your work with the database. But one thing that I

noticed in the guideline values was that there's a bis-phthalate number for the 10 day *Hyaella* mortality of 500 and then there's a guideline for the 28-day *Hyaella* mortality which is orders of magnitude higher. And to me that's counter intuitive. One would think it should be more sensitive. I think this particular chemical is going to be one of the challenges that we're facing as we're moving forward. So I'm just curious as to why there's such a big discrepancy in those two numbers from a *Hyaella* mortality test configurations. Hopefully that's not too detailed and technical for this.

TERESA MICHELSEN If I had to guess, it would be that they are from totally different data sets. And this points out there are going to be occasional issues in the data set like that. It should be that phthalates have, are one of those co-varying sets of chemicals. We haven't tried to add them together or sum them up, but I have noticed that before with the phthalates that there are some. I call it a seesaw effect, one goes down and then the other one goes up. And when the other one is down, the other one is up. So that tends to happen with co-varying chemicals. So there could be some phthalate co-variance issues in there. It also could just be that the long term test is from a totally different data set than the short term test. And that one's from more eastside and one more Westside, and one is more urban and one is more rural. And I don't really know that that's the case. I'd have to actually go look at the data sets. I think someone else may have noticed that particular issue. So there are some individual numbers we can look at. If anything looks odd, there is the possibility that it is off for some reason and we may need to look at it again. Whenever you have a multivariate model that has co-varying chemicals in it and multiple solutions, you have to apply the common sense test to it. So, we have done that and the numbers we have in here, most of them look very, very good. They make sense intuitively and even reasonable, but if there are ones that jump out at you as being strange, then we need to look at that. We'll go ahead and make that comment and look into it.

LAWRENCE MCCRONE (Exponent) I just had one point of clarification I was hoping to get an answer on. Teresa said something to the effect that Ecology never makes decisions on the basis of a single sample. That it always takes a minimum of three samples to make a decision. And back in the days of promulgation of sediment management standards, was my interpretation as well. But more recently I've heard other Ecology staff say, not anyone in the room today, say that Ecology can make decisions on the basis of single samples. And I just wondered, given that we have some other sediment folks from Ecology here, what their interpretation of that is.

CHANCE ASHER The interpretation of the rule is as you and Teresa both stated. In order to list a site, we need have to have at least three samples and has to meet specific criteria. While I don't know the specifics of whom you've talked to about this, anyone who is saying they can base a cleanup decision on one sample is mistaken.

LAWRENCE MCCRONE It's the difference between listing a site and making a decision about the need for cleanup within a site.

CHANCE ASHER No, that would be, the decision to list a site would be a decision to require cleanup. In the rule that's the way it's written, that it would require at least three stations.

LAWRENCE MCCRONE But there are large sites, like the Lower Duwamish, for example, where there are isolated SQS exceedances for example, and surrounded by many other stations

with no SQS exceedances and what we've been told is that Ecology can require a cleanup at that location based on that SQS exceedance.

CHANCE ASHER Well, we can't answer site-specific issues. I don't know the rest of that discussion. I can just state that in order to require cleanup, and to list a site, we need to have three stations, the average needs to exceed the CSL, as well as other requirements that need to be met.

LAURA INOUYE I think the answer to that one is that whole area is the site, so where the individual cleanups go, because it's already 'a site' that that decision that that's an area that needs to be cleaned up, that doesn't need three because it's already part of the site for this discussion.

CHANCE ASHER I would hesitate to comment any further because I don't know the specifics about the Lower Duwamish and those decisions. But I can tell you the requirements in the rule are very clear.

ROB PASTOROK Just one more on the multiple solutions issue. How does that arise and does it matter, like are the chemicals entered in an order, and does it matter what order you entered them?

TERESA MICHELSEN They are processed in a specific order. The model starts with the one that has the most false positives. And it reduces it to the point where it equals another chemical and it moves to that other chemical so it kind of brings the false positives all down consistently moving back and forth among the different chemicals and that over ten years or so I refined that algorithm and that seems to be the one that produces the best reliability. But yes, it makes a difference sometimes where you start the model and what result you get. Which is one of the reasons I run it multiple times over and over searching for that solution that gives the best reliability. And also looking for individual chemicals or groups of chemicals that seem to be co-varying and also running co-variance analysis to look at that same issue. Then there's a final step where we'll take the lowest value of all those that we see and work up. So that tends to counter the potential tendency to pick too high of a value because some other chemical is holding down toxicities. So we'll start with all the lowest values and work up and that gets to an even better reliability. So a lot of this has been trial and error trying to figure out how to address co-variance in a reproducible, programmable method, which everyone insisted on. And it works pretty well, but it's not perfect.

CHANCE ASHER And so for the next hour and 20 minutes we will be talking about the human health background issue, and the accompanying summary document which was sent out to everybody. And we'll take comments on this human health/background summary document. Thank you very much for the comments received and to the folks that didn't send comments, that's fine, we'll hear what you have to say during the meeting or afterward. What I could do is go through and clarify a few things based on some of the comments.

So just to clarify, one comment from Pete Rude was for clarification about what this document was. He asked about the introduction to the summary "to get some concrete recommendations" from the group. He was just wondering what that meant if we were looking for group recommendations or consensus. What we were attempting to do was respond to what we heard

from the advisory groups who have asked us what we are actually doing with what you are telling us. And those are great comments, good feedback, we appreciate that.

It kind of took us a while to get to the point of trying to show you what we're doing because we had to really work through some of these issues and think about them. This document is an attempt to synthesize what we've heard, melding in some of the things we've already proposed, to see whether or not that's a framework we should continue discussing. So what we're trying to do is incorporate some of the issues that have basically floated to the top or the consistent concepts or the ideas that we've heard from you that may be feasible, and modify our initial framework that we put out there for you to think about and basically let us know what you thought. And that's the idea here and we're not looking necessarily for consensus from this group. We're not looking necessarily for group recommendations. You can certainly give us your ideas as individuals or as representatives of your constituencies separate from what is in this document. But that's basically the gist of this summary document. It's an attempt to try and get us moving forward rather than continuing to discuss the original options Ecology proposed, just listening and then Ecology going back and talking about it and trying to figure things out. So I'm just putting this out there to see what you guys thought. So it's an attempt to try to show that we have been listening and we've only got a few meetings left so I want to make sure that we're hearing you. Based on some of the comments that we've already received from you, this seems to be working. And again this summary document wasn't meant to capture everything that we've heard. It was meant to capture consistent, feasible concepts that we've heard so we may be able to move forward. This does not mean that Ecology's made a decision at all. We've still got a lot of internal discussion to do. But this may give us a framework to continue moving forward so we can finish these meetings and then go off on our merry way and have our internal discussions and come up with some ideas that will work and even some rule language at some point. So, that's addressing one bullet of Pete's e-mail.

I originally would have preferred synthesizing your early feedback on the summary document for you but there wasn't time. We appreciate the comments coming in no matter what. But we can still go through your individual comments. Perhaps what you can do is summarize in bullet form, some of the bigger issues that you have and I like to just discuss those.

TERESA MICHELSEN I have a request. I'd like to make sure most people have probably read them but for those of us who sent them out to first have a chance to express any thoughts that they have.

JOANNE SNARSKI I know I didn't get any out formally because frankly I didn't have the time to formalize any thoughts. I think it was a good first attempt. I think my first impression of it was that it didn't jump out and shock me in any way it seemed to capture a lot of the information. Actually what did surprise me is that comments that came back on it. I haven't had a chance to go over all of them, but some of them I didn't understand.

GLEN ST AMANT I think what I'd like to do is weave mine in as we go. I had some similar ones to ones that were articulated for efficiency we can weave through.

CHANCE ASHER Okay that's good. So why don't we start out with Pete since I already started out with Pete's comments. Pete would you like to go ahead and synthesize in a little less detailed form your main issues and we can work through those.

PETE RUDE Yes. I think one of the things that I have been thinking about are all the things that we've been kicking around. Looking at the summary you put together, I'm getting overwhelmed in my own mind about the technical challenge of some of these. Do folks have some ideas about short cuts?

There's some things that conceptually make sense that we can say, oh yes that's logical. If we can decide what sources come from the PLP and make sense of those. How do you, on the ground, get some of that done in a reasonable amount of time so that was one theme. Another one was, I'm still struggling with this idea of regional background, and again it's just sensing, and I love to hear comments from others, that it's seems it's just so big of an area for numbers, and maybe that's what they need to be, but the numbers just might not help and moving things forward I just saw potential for grinding back to some of the paralysis we've seen on some sites. And so I think these are two of the main themes. I had a few other specific items that I put in there, but I think I'll stop there with those themes. I hope those kind of shine through with what I wrote.

CHANCE ASHER I think they did. What I like to do is open this up to you guys to have Ecology speak less and you speak more. If you have any responses to Pete, or some possible solutions or even steps forward for Ecology to address, or do we need to continue on with some of the specifics in order to get there.

CLAY PATMONT I don't know if you had a chance to look at what Teresa wrote, but I think some of the approach of how you would define regional background. I agree that it looks like it's -- depending on how you configure it -- is a very monumental task that we can never get two people to agree on how it would be done. We just need to be more specific as a group on those recommendations. That's what I took away from your comment before. We were presented an approach and a number of concerns about that approach. Regional background is a concept that seems like it holds promise but the devil is in the details so let's just try to actually explore it a little bit more. I tried to break it down a little more going off of what you put together. Teresa, in terms of a five or six step process eventually defining it.

Bullet five: I stole this mostly from Teresa. I thought it was important to come up with a definition. I know that we've bounced around "hydrodynamically influenced." Personally that doesn't do a lot for me because how many sites do we actually know the hydrodynamics of the site well enough to do it. I think that's a great one. It sounds really good until you try to actually define it and then it can be defined so differently depending on the methodology used. So first you would say it's really kind of a watershed or embayment scale. But it not influenced by releases from specific facilities so I think there's a complexity that is huge. We want to come up with a system that will work - that will be simple if all you need is a simple solution but it can go complex if that's really what you needed to do. Trying to build that into the definition that regional background cannot be influenced by a group of facilities that you've already identified that definitely includes combustion and regional storm water sources. That's the first point that I would think of.

CHANCE ASHER A clarification. One of my notes that I had on here, when you say a specific facility do you mean a PLP source?

CLAY PATMONT I guess what I would regard as an upland site, the way it is commonly defined in MTCA. You got a release you've got a property or properties that are affected.

PETE ADOLPHSON I think I understand the overall direction you're moving but then we start going into things like influenced, not influenced by and then it starts to get a little more fine so we need to understand that.

LON KISSINGER If you look at PCB concentrations in sediments in inner Elliott Bay rather than outer Elliott Bay. Clearly there's been influences from point sources, perhaps from PCBs from Boeing plant and that flowed out and so, but you can't really explicitly tie in. You don't see and neatly tie in a contaminants transect and watch it drop. So maybe that's it, as you can't deal with the diffuse PCBs in inner Elliott Bay, or maybe you could.

TERESA MICHELSEN There's two points I like to make with regards to these definitions without doing actual calculations. One is, one thing that I saw in Pete's comments as a real justified fear and concern about how we gather this data in a reasonable timeframe. But I want to go back to the BOLD example and back to inter-agency programs and coordination. In the Puget Sound region, when we put our minds to a task and we get all the agencies together, we can do amazing things. That BOLD study was done really fast with all the background data which is going to serve a whole lot of different programs. That is just one example of the last 20 years of projects. And I go back to the original urban bay action teams. We need a reinvigoration of the urban bay action teams. We need to go in on a cross-program basis and look at each of the major urban bays and do some data collection. Get refocused on that for a few years, get some Puget Sound Partnership money, whatever we need to get focused on grant funds and program funds to do that. If we put our minds to it I believe that we can get the data collected in a reasonable timeframe. We might have to go bay by bay, we might not have it all at once but that's as fast as we can make decisions anyway. When we collect data in Elliott Bay it's going to take us a year to figure out the regional background with that data anyhow.

CHANCE ASHER Clarification. You're talking about data collection to calculate a regional background or natural background.

TERESA MICHELSEN What I'm trying to say is to regain that cross program/cross agency spirit of data collection and not feel like we can't do it because we're trying to do it within one specific program at Ecology; that would be next to impossible and it would take forever. But I think that it's proven as a program over last 20 years we can take on these tasks and we can carry on this kind of data collection alone so that's just one part of it. The other part is don't be afraid to use professional judgment. It's going to be a judgment call for a lot of the stuff. We're going to have to look at the Elliott Bay data and say: There is this diffused range of PCBs out in the bay. It has something to do with where that drop-off is. You can kind of see the site signatures of that drop-off. It's this range of fine particles. So maybe we should define it at the drop-off. We first define sites in Elliott Bay and put out our site list. In early 90s we did it just this way. We got all our GIS data, we sat around in this room, and we said what makes sense. And all of the sites overlapped in Elliott Bay just as they do now. It's not different now from what we did then. It's just that not being afraid to do it and not being afraid to put your rationalization out there to the public and say this is how we decided it. It's not going to be perfect. If we just sit around and say it's too hard, it will never get anywhere from where it is now and right now we're stuck.

CHANCE ASHER Are you saying that we need to get that data and establish some numbers before we establish a rule or promulgate a rule?

TERESA MICHELSEN It depends upon what the rule says. I think we need the regional background information. We may need it before we start holding people to it but the rule could incorporate that. You could have a case for when a regional background level has been established and you can have a case where that hasn't been done yet. Maybe it needs to be done more site-specific in cases where it hasn't been done yet. In places where it hasn't been done yet, you have to deal with it at the really big sites that have the resources to try to figure out what regional background is.

PETE ADOLPHSON I just have a follow-up question about that proposal. Do you think that we would need to define what "regional background" is before we go out and collect data so we know what we're doing with that data in terms of our evaluation.

TERESA MICHELSEN You mean in terms of decision criteria?

PETE ADOLPHSON Right.

TERESA MICHELSEN Ideally, we've done that with the BOLD data survey set and that's why I feel we can do it. We wrote up what questions we were trying to answer with the question in the BOLD survey in the sampling plan and specifically how we were going to use the data to answer those questions before any data were collected. And then we applied those decision rules. We made a few decisions by station corrections as the situation called for them. We followed our steps. And that's why I feel like I could do it that way we can do it again. It's just a matter of being bold.

CLAY PATMONT One concept that I have. I don't think anything we are doing here is exact or hypothesis testing or anything like that. This is maybe an overlay of my comments, but I'm writing this from the standpoint of how do you incentivize a willing PLP who really wants to get something done and doesn't want to get caught in the mud. I think actually, with the definition of regional background, a reasonable agency and a willing PLP might not agree as to the basis for how you're defining regional background, but I would bet you they would agree with the physical dimensions of what we need to deal with. They might have entirely different reasons for why we're coming at.

CHANCE ASHER You talk about physical dimensions, do you mean you mean geographic boundary?

CLAY PATMONT Like what part of Elliott Bay would this PLP actually address. There are just so many different ways to answer that question and I think when we get together in the joint meeting we will hear more of that. That's my bias with this whole thing. None of this is amenable to detailed scientific testing. But I do think that there are reasonable people who want to accomplish cleanup. So from the PLP side, they just want to deal with this to get it off their books. They want finality and they want a covenant not to sue. I know this doesn't come free, not everybody is going to be motivated that way. So that's where my head is at, to make these rule revisions work for cooperating willing parties. You got to deal with the other side -- what about the parties -- I totally understand -- that needs to be addressed. That's not -- that's not what I'd

like to contribute to this whole thing. I'm trying to make this system when we've got Superfund, a great giant hammer out there, it's really used for. So, that's an overlay. So then, coming up with a definition that could be simple. We have a lot of bays now, with a ton of data. Six more now.

TERESA MICHELSEN If we get back to Pete's comment. We establish regional background not in the context of that site RI/FS but as a separate Ecology activity like we did with the Urban Bay Action Team report where we said "What is the state of the Bay?" with respect to source control. It will be taken out of that polarized antagonistic context of a given site cleanup. This is our understanding of the bay, background and regional background. We're going to start using this. That's a better approach than picking the first big nasty cleanup site that comes along and trying to hang it on that.

CHANCE ASHER Glenn, did you have a comment?

GLEN ST AMANT No. I agree.

LON KISSINGER This was based on the statistician who had a workshop on dioxin. If we want to use that as a criterion, I think we should bring that in and just discuss it as a group too.

TERESA MICHELSEN Yes I'd like to have another statistics discussion too at some point.

CLAY PATMONT Basically the process is to first identify the natural background and I think the BOLD data set is good. And so you define what that number is, you go into your embayment and you identify where at the margins of the bay, you're in that sort of natural background, and you take that out of the pool.

CHANCE ASHER How about the comments that we've heard about the BOLD data being used for the background in some areas outside of Puget Sound proper. Those concentrations at reference sites, theoretically non impacted areas, those concentrations may be lower than the BOLD data. How can we incorporate that?

TERESA MICHELSEN They are lower in the Strait of Juan de Fuca and the coast, they are lower. You've got to have more data. I think we have to start with our most contaminated areas for our Puget Sound urban bays and for other areas it's going to be harder.

CHANCE ASHER The natural background data would apply to certain parts of Puget Sound but not necessarily those that are outside of that influence of urban bays.

TERESA MICHELSEN It applies within the area where the data was collected which doesn't include the outer parts of the Strait and the coast.

LAURA INOUE There were only a few and remember the main reason the DMMP group who was the one that was trying to get the BOLD initially to take a look at the dioxins and then it spread to all the other contaminants. When we were looking at gathering data, it was in areas around while it would cover the places we were doing our disposal. In the Straits of Juan de Fuca we had a few but we stop fairly far in, because we don't have disposal sites out there so why would we be characterizing up more towards the open ocean? So is that a problem with the BOLD data that tend to be more on the siltier side and closer to the main urbanized areas than

way out in the Straits. The other problem with the Straits is the worst days of collection were just outside of Admiralty Inlet going up towards Rainier. You get out away from the shoreline all you have is rock and grab after grab we had rock.

CHANCE ASHER So let's go beyond the specifics of BOLD. Do we need to go through the rest of your points, Clay so that people understand what you're trying to say?

CLAY PATMONT Yes, this is where I was trying to go. I think the BOLD data is going to be worthwhile for using in 90% of the situations. There might be some where it's not but this is the place to start. The third piece is where we are right now and this gets into this quality of how do you determine what's in the influence of a site versus the smear that's out there? But there's going to have to be a value judgment looking at footprints of chemicals -- there's a bunch of tools. But going back to Pete's comments, you start where a simple tool will suffice and with a more complicated one, if there's a reason to take it into more detail, you would. So you might start relatively conservatively.

CHANCE ASHER We like to get some more comments from folks about this layout that Clay has done because it's so specific, not necessarily that I want to give Clay extra time or anything. It's very specific so go ahead.

LON KISSINGER One question, why would you pull out areas below natural background from within the bay.

TERESA MICHELSEN Outer boundary. You need to include this because in that area you need natural background if you have a site in that area.

LON KISSINGER So wouldn't look at the shoreline, you're looking at the outer limits of Elliott Bay, once you get to natural background, that's the end of the embayment. But within Elliott Bay, if you had an area that was below natural background you would keep it in, I would think, because otherwise you would be biasing your regional background high.

CLAY PATMONT Unless it's associated with some really unique feature of the bay that is behaving differently but yes I think the purpose was mainly to draw a broad outer boundary.

TERESA MICHELSEN Part of my reason for wanting to take it out was, in Commencement Bay we already know a lot of Commencement Bay meets natural background. So my point here is if you have a site within the area that already meets natural background, the natural background is your cleanup standard not the regional background. It's probably good to draw boundaries within which regional background applies rather than natural background applies. I think it's important not to apply a regional background that's not already as clean as natural background and we already have that in urban bays are parts of urban bays that are pretty clean. We just need to know where that boundary is and then we can decide to include any of those data points in the regional background data set.

LON KISSINGER Might be helpful to actually just look at the data for particular urban bays. We already know what we're talking about with Elliott Bay and Commencement Bay.

TERESA MICHELSEN We just need to put it on a piece of paper. Part of it is, maybe we're just not looking for the particular statistic. If you call it the 90 percentile then you have to worry about whether the natural background part of the bay's in or out but if what you are doing is looking for a breakpoint between the smear and where the sites start to become obvious you could be looking for a breakpoint instead of a percentile.

LON KISSINGER Sounds like you're coming up with a different proposal.

TERESA MICHELSEN I didn't have that as a new proposal. I just am thinking out loud. But some of it is just mapping the data and getting out there, and looking at it as reasonable. And that's how I think we define a site in the first place. Probably what we have to do now is to see if it's as complicated as we're making it or maybe be obvious where to draw the lines

CLAY PATMONT The way I look at this is, you look at it like the Elliott Bay example is the most complicated so we have to understand Elliott Bay and then choose maybe a simpler bay. Maybe some of the sites up by Everett, Bellingham, Port Angeles or whatever are pretty complicated. They have a number of sources but they're probably an order of magnitude, less complicated. And then there's the relatively simple site. I think there are a lot of those third category sites out there and those are probably the ones where a PLP is very happy to go to because there is nobody else around. Their job is pretty clear. You don't want to have to drag those guys through the entire Elliott Bay process. That's the way I look at it. And, going through this first cut, you might start out by saying all of the Elliott Bay or just a chunk of Elliott Bay or something like that. But, other sites you might just want to have a reasonable management unit you can pull right into the program. That's where I was going. I was thinking that there is a much more important step to Elliott Bay and that's this weak contamination. Sort of a potential idea which is something difficult to quantify. But with our best professional judgment and tools that we have, if you have a group of willing parties you might be able to get there. So, is that clear? Basically the regional data is defined as a preliminary cleanup level that you might want to think about hauling into the program. I was struggling with Ecology's point of view. Is Ecology wedded to the point of view? Let's say, that if you define this, this way, then the entire bay all of Elliott Bay is a site. Maybe it's not really a technical issue, but from a program issue, if you made every PLP in Elliott Bay have to work together it's a recipe for disaster. So this would be something that might work for some sites, but, with Elliott Bay has to go to the next level.

CHANCE ASHER That was the concept behind what I've heard from you guys and also from other groups that you're just going to grind things to halt if you're going to have an entire embayment called a site and then you bring in all the PLPs. Divide it up into some manageable chunks and that's the unit concept that's somewhat similar to the superfund operable unit idea. I think we can make it better. But, that's what we heard from you guys, and then kind of separating that out into short-term goals, actually getting those sites that have a high level -- relatively high level -- of contamination compared with the rest of the bay, getting them cleaned up and resolve some liability for the PLPs. We've still got the rest of that embayment that by law is still above natural background that could be considered a site. We can deal with that in a different way with that long-term conservative goal of natural background whatever we decide that is. So that is basically what we heard from you.

TERESA MICHELSEN One of my main comments is not to forget about tools we already have that were used in the past and we've never tried to define Elliott Bay as one site even though,

never mind bioaccumulation. Just because all of Elliott Bay exceeds our cleanup level doesn't mean that it's all one site. The rule defines a site as exceeding a certain level, yes, but also having a specific chemical signature. That's in the rule. So they are not site units. A site unit is intended to be an engineering term for how the site is cleaned up. It is a separate site. So, PSR is separate from Harbor Island, which is separate from Lockheed shipyard. They are all separate sites. So I would stick with that paradigm. Just because we're using a different cleanup level for bioaccumulation, let's keep the same basic rule structure that says all of Elliott Bay exceeds these bioaccumulation guidelines in Elliott Bay. We have 10 sites and they're defined based on their sources and their chemical signatures, and they're all adjacent to each other perhaps and they've got boundaries that are administratively defined based on chemistry or property boundaries even in some cases, where it's convenient to do so. That's what the PLPs want to do, but don't call them all one site. That leads to huge messes.

PETE ADOLPHSON I don't know if that's, in terms of site definition, if that's what Chance is referring to.

TERESA MICHELSEN That's the way the terminology came out in the paper that we got and I was quite concerned about it.

PETE ADOLPHSON I think the concept is more that there are separate sites that when you go beyond where you can distinguish them as separate sites they have a mixed plume that is beyond cleanup levels. That is where it becomes difficult. It is the mixed plume. But, that is where the difficulty comes in. How do you address that because you have multiple PLP's that are responsible for that mixed plume. Everything is on the table at that point. How do you address that?

TERESA MICHELSEN So let's start with that. You might even define that as a separate site. Just because you clean up of all these sites and one would be left. How do you handle that site or how does each PLP resolve their liability for that site? For example: with buying into a fund. So that they get a final buy off and that big site out there, which is just a big smear, gets monitored over time by Ecology using that fund. That was my thought, about how we can try and deal with the actual liability those people have, but also give them their buy off. You can easily calculate the net present value of that monitoring cost and include that in the sum.

CHANCE ASHER We tried to capture that in the write up of the summary or whatever terminology we decide on. It's the same concept you're talking about. The individual sites or units, or whatever you want to call them, that a PLP is definitely on the hook for, defined by their chemical signature. As Pete said, with that mixed plume, whether we call it a separate site or not it doesn't matter, the same issues are there. How do you deal with getting them to cleanup those areas that we know they contributed to? The concentrations are going to be a lot higher because they're closer to shore versus the mixed plume that is still above background concentrations. And it's captured in the text as well. What you brought up and what we've heard from other folks is to give them some finality for what they actually, physically clean up. But you can easily say what they're on the hook for and that concept of a communal fund has been brought up numerous times by numerous people. Look at that larger site and the mixed plume area as monitored or natural recovery or what everyone calls it. You put dollars into a fund and verify that actually is recovering.

TERESA MICHELSEN The reason we keep having this fight about the separate sites is that, even though the concepts are the same, it's really important from a regulatory standpoint. It's had the same problem as sites being pulled it together and kind of a negotiated solution. That would be horrible. And also I didn't get a clear sense that Ecology was willing to give them a buy off in pieces in any way.

CHANCE ASHER But what we're trying to do, is show you what we think we've heard and not what Ecology is willing to do. At this point, I'm just clarifying what we're hearing from you so we can put together a framework, which says generally this is what we've heard from our advisory groups.

TERESA MICHELSEN They really have totally different regulatory implications.

CLAY PATMONT So I'm going to try to take the unit concept one step further, like in this point 8. I'm not sure what the optimal balance is between units and sites, but in these complicated situations, try to identify units that make sense to remediate. I don't know if there's a downside to Ecology allowing a PLP to say "make us a proposal, " come to us with what makes sense and you only get that covenant not to sue for what you actually address. There are ways to do it. This is where it gets a little dicey in terms of the re-contamination potential. Source control or recontamination is one of those things that looks like it's an easy task until you try to do it and then you realize it's really complicated. Even when you're sampling and doing a lot of modeling you don't have any better of an idea of what's going on. You open your eyes enough to see that it's a lot more complicated than you first started. That being the case, there are still tools. And the best professional judgment, to overlay. And I think the huge one is the overlay of storm water sources. That's what I see and trying to make a reasonable apportionment of the data you have, you might decide that you have too much uncertainty to deal with the next-door neighbor. You have to fold it in your thinking, but you still have a hotspot. No, I don't like that term either but we've got to come up with -- instead of the schmutz -- these are low concentrations by everyone's standards but they're above natural and regional background. You try but you might not agree. Ecology might not agree along these lines but they just might agree. And then to the extent that this makes sense to address now, you're held to that sort of re-contamination potential standard. And that's possibly going to be more complicated but I don't see any way out of it, myself.

LON KISSINGER One thing I thought about was trust funds or something like that. We're talking about sequencing remediation of the sites to address re-contamination issues. That's the only disadvantage I see to just sort of plowing ahead, instead of just doing it independently and not looking at the larger system. I don't know whether that can be estimated in such a way that sufficient funds can be allocated or a commitment can be made. I guess it's like re-contamination. Does it make sense to take care of the worst first so that they don't slough off and hit a less contaminated site and then you do a sequencing of worst to least contaminated but still needing to be addressed?

GLEN ST AMANT So, Lon, would that be like anticipating having to go back and do an enhanced natural recovery.

LON KISSINGER That's the problem. You really know how much you spent if you actually spend the money, right.

TERESA MICHELSEN These are societal issues. Thinking about the waterfront issue, I was really involved in doing a big survey of all the sources of contamination on the waterfront for NRDA. We were trying to decide which areas we should go in and spend money to restore and cleanup, thinking about which would get re-contaminated and which wouldn't. We have the luxury of choosing where to put our money, which is a little different when a PLP needs to resolve their liability. Let's say you have a big industrial site that's contaminated and it's a mess. Everyone agrees there's a big legacy problem that needs to be resolved. And maybe right in the middle of the site, there is a King County outfall that will recontaminate with phthalates.

We'll use an existing scenario. I honestly don't see why the PLP should even have to worry about this City of Seattle outfall. I think the PLP should just have to clean up their own historic contamination, deal with legacy contaminants. And yes, we know, that the outfall is in the 50 year plan. Seattle has their plans on how long it's going to take to control discharges. They've got a 50 year plan to do that; they're going to have to clean up whatever comes out of their outfalls. That's the way I think it should work. It's not that the original PLP's issue nor should they even have to do a study of City of Seattle's outfall and recontamination potential. That's my opinion. I think it would be easiest if we'd recognize that there are going to be some legacy issues, like the twin 96ers in Commencement Bay.

LON KISSINGER We talked about this before - kind of changing the public perception. We were going to get rid of PCBs, which are legacy contaminants and for ongoing contaminants like phthalates and PAHs. The agencies don't want to get beat up in the media by saying: You you blew all this money on clean up and look, it's re-contaminated again. What are you doing?

CHANCE ASHER We need to figure out how to address that public perception issue. One response Ecology has to the recontamination issue is: We cleaned yo a very high risk area to humans and the environment from highly toxic chemicals at extremely high concentration. Even though the sediments are recontaminated, the contamination is different chemicals at very low levels. So, the reduction in risk accomplished in the cleanup is very beneficial. Joanne, I'm curious if we can speak about some of the issues the ports have.Or if they would be supportive of what we've been talking about with providing some final settlements for those areas that we can define by chemical signature or high concentration areas from that PLP even though that PLP is responsible for some of that bay wide contamination. To separate that issue of contamination from the areas or sites, whatever terminology we end up using. From the ports' perspective with all the dredging issues that they have, if that would be something they would be supportive of? Or if there are some specific issues that we haven't thought about.

JOANNE SNARSKI Yes, because of maintenance dredging.

CHANCE ASHER Let's take the Port of Olympia, let's use that as an example. So, there was a dredge done. We knew that those dioxin concentrations had contributed to the whole inlet contamination.

JOANNE SNARSKI We knew that would re-contaminate.

CHANCE ASHER We knew that would be re-contaminated but let's say in that scenario if there was a way to actually prevent the re-contamination from the PLP let's say it's the Port. Let's say

we could somehow prevent that by taking care of the whole area, let's say we provide some type of settlement for that area that can be feasibly remediated. Yet that PLP, because of what was either in the sediment in that area or a discharge that they're responsible for, has also contaminated 7 miles up the inlet with dioxin. We can somehow separate out those two issues to get some finality for what the Port or another PLP can feasibly remediate and then separate out the long-term or communal schmutz that they also contributed to. When you're always doing dredging, it seems like - otherwise how would this work if we had all of Budd Inlet as one site and insisted it's all dealt with.

TERESA MICHELSEN For most Port districts, except for those areas where ports are going to be responsible for that that was Cascade Pole caused the problem. The ports only responsible to do its dredging and potentially spreading the material. Let's say that during the dredging some material escapes.

CHANCE ASHER We've had ports as PLP's. That's the reality of it. I'm not saying that the Port of Olympia was a PLP.

JOANNE SNARSKI Well, it does highlight more of the complexity of the issue, in particular in our case we indemnified Cascade Pole. We can't even go after them. So working with them to make that all come together is how that's going to have to move forward. But you know, you have to look at all the other sources. I thought a lot about the definition that you put out here. It's very interesting from the dioxin perspective because of combustion and storm water and that sort of thing. How do you move forward and take a phased approach? I can't say that I have the answer to that but that's certainly the way we have to look at it.

CHANCE ASHER Do you think that the ports would be supportive of us trying to figure out a way to try and move forward?

JOANNE SNARSKI Oh absolutely. I think these concepts came across in the summary framework, that was out there you know. It's kind of the short term goals and the long term goals. You know the Ports are here to stay. You know we are not going anywhere. I think that's what differentiates Ports, from let's say small land owners. You can bankrupt them really quick. You know Ports have different resources and more resource to more public grants and what not, if those continue to stay around. So Ports are part of the community. They are there to support environmental objectives and all of that kind of stuff. So absolutely. But the key is how do you get it phased in and try to meet that working waterfront need? I felt the concepts were thrown in here but the devil is in the details, absolutely. For me that's where it goes.

CLAY PATMONT I think we have talked about it before but I just want to mention it again the role of the Clean Water Act and other programs to deal with the longer term source control. I just don't think that MTCA/SMS is designed to deal well with source control and on a larger scale.

CHANCE ASHER Can I clarify for the record do you mean source control as in stormwater?

CLAY PATMONT Yes, I'm sorry. Yes absent the basic facility and the specific issues you have, soil issues to take care of and a groundwater issue, those kinds of things. And again MTCA does a good job of that. But the problems that we're dealing with, the more we learn, it does seem

to be on-going storm water sources from a very large area. I wasn't clear on how you were summarizing our recommendations of what you were hearing. But I think I just want to be really clear. I really think on dealing with regional background data and natural background is just way outside the domain of what SMS or MTCA could ever deal with.

CHANCE ASHER So do you mean, to clarify, dealing with how to go about decreasing concentrations from a regional background concentration down to a natural background on a bay wide scale? The side of the feasibility part of both rules and its a better todeal with under the clean water act as a source control issue?

CLAY PATMONT That's the way I feel. It's really a societal issue more than a PLP issue.

CHANCE ASHER Does anybody else have any specific comments about that? The SMS/MTCA rules and the ability to deal with source control/stormwater, getting concentrations down to natural background..

TERESA MICHELSEN I think I made the same the same comment, in the sense that I think that clean up should focus on legacy contamination and stormwater source control should focus on the remaining reductions in contamination over large areas. But I don't agree that SMS wasn't designed to do it because SMS has both source control and clean up in it. I just think that the SMS authorities aren't used within SMS the way that they used to be. So their use is largely diminished to zero from what I'm hearing. And needs to be reinvigorated. I hope that the MTCA folks and the TCP folks will consider and not take this all the way down to the natural background within MTCA. I think that is just not what MTCA was designed to do. I mean urban levels of contamination exist on our streets and in our soils. You know it's a very similar issue. But MTCA doesn't go after all of those streets as contaminated sites. You need to apply some basic common sense here. Use MTCA for the same legacy sites down to regional background. And use our source control authorities below that. That was my big comment.

GLEN ST AMANT I don't feel well-versed enough to answer the question of what the Clean Water Act authority should be for source control versus SMS and MTCA cleanup authority. We talk about your hot spot or your warm spot clean up, but then the next phase and I think that really is the rub for the next phase. Don't look at me for guidance on whether it is Clean Water Act or MTCA under the SMS.

CHANCE ASHER Ok how about let's ask in a different way. The next phase to deal with our long term goals for an entire embayment to get that back down to natural background would that be a cleanup issue? Do you think we feasibly could obtain that by requiring a clean up or could that be obtained by addressing source control on incoming sources.

GLEN ST AMANT So what I guess I have a hard time discerning is there going to be ongoing discharges from a site?

CHANCE ASHER No I don't think so. We are really talking municipal storm water. Not sources from a site or a PLP.

GLEN ST AMANT I guess I'm talking storm water from a site.

TERESA MICHELSEN Atmosphere and deposition.

CHANCE ASHER What Glenn was saying, was storm water from a site, a PLP.

GLEN ST AMANT And to the extent that SMS needed to regulate that to an appropriate level.

TERESA MICHELSEN That could get done through our Water Quality authorities. Or MTCA either one could be used for a PLPs' stormwater.

CLAY PATMONT Yes, that's what I was thinking too. To the extent that it is under the influence of a facility/PLP, you could actually do something about it under MTCA. That would be part of the covenant not to sue. What the standard is, you know you've got atmospheric input that is going to take it to a certain level. I think that's where the flexibility it's going to be site specific. If you take care of the legacy issue, but then say ok, you've got atmospheric or other watershed sources, even if you clean it down below regional background, it would readily come back up to that level of recontamination. You make a site specific determination; you've got to control your sources down to a reasonably attainable level but then there is all of this other stuff/stormwater which is out of your control.

TERESA MICHELSEN Make a common sense decision not to include that part ever in MTCA, otherwise all of this stuff gets tied into the liability and it's not the right tool for this. We have other regulatory authorities that are a better tool for that. I will say is being super careful with the language. Because we have a seen what the last 20 years we've made with some of the choices in the language have come back to haunt us.

CHANCE ASHER Is there any specific language other than the word 'site'?

TERESA MICHELSEN I made a bunch of recommendations in the comments that I have made that I don't really want to go through. But every time I saw something in your write up that I thought that was something that we needed to be careful of I did note it. I will look at the new terminology things that you've wrote up the definitions. I think we just really need to be careful.

CHANCE ASHER I'd like to give Lon some time he put together some pretty extensive coments as well. To talk about what some of the main points were. And see if we can work through them.

LON KISSINGER Well, I was just kind of wondering, per the last conversation, whether we might want to talk to the people working the TMDL programs. I mean they might give us some tools or thoughts or concepts. I'm not with TMDL.

TERESA MICHELSEN Yes, TMDL has been used in Portland Harbor and in Ward Cove. It's been used in other places in Region 10. But for some reason it has not been used in Washington. I think that is an additional tool that we can use.

CHANCE ASHER Specifically, what nuggets would you pull out of the TMDL program?

LON KISSINGER For TMDLs, you're looking at a body of water and you're looking at all of the sources that have input with that body of water. You are allocating loads to them and I was

thinking we're kind of doing some of the same analysis. I'm not really familiar with TMDLs but I'm just thinking after listening to this conversation that they might have some ideas, thoughts, processes that might be helpful in our deliberations.

TERESA MICHELSEN Recently, they have taken counting the sediments as a source. So it's a way of taking the smear that's left over and saying: how do we make sure this doesn't get any worse? And furthermore, how can we help it get better? What source controls do we need to produce these levels? I know I brought up TMDLS before. I do think that it's a tool that is being used in other parts of Region 10 and can be used here too.

LON KISSINGER As long as you keep the concept behind bars and just look at it. I guess my major comment was: ok we've decided that a lot of these sites have risk based concentrations that are below background and we are going to deal with them on a background basis. I think we need to dot that (I) and cross that (T) and do something to look at that. So I know that Teresa actually made that comment. So think we just need to actually do a comparison between risk-based concentrations and background. We can either do it on a tissue bass or we can do it on a sediment basis. There was work done in the mid to late 1990s that might have some relevance there. And I do think we need to decide what contaminants we are going to look at. I would say most likely for PCBs, dioxins/furans that indeed risk based concentrations are below background. Are there any more exotic contaminants that aren't used as broadly with limited distribution and the risk based concentrations truly are above background where there really isn't a relevant background? I don't know; are there any major sites where there were pesticides or herbicides used or arsenic. It isn't really bioaccumulative. In a limited sense, it doesn't tend to jump trophic levels.

However, we have direct contact risks that are above. I mean the risk based concentration to get acceptable or direct contact risk is below background. So it doesn't even have to be bioaccumulative.

So, anyway, just justify it. And the State Department of Health while developing its seafood consumption advisory for Puget Sound did do some of this analysis. They really dropped it down to PCBs and Mercury. So that might be a place to start.

TERESA MICHELSEN I wouldn't rely on a Department of Health analysis.

LON KISSINGER No, but I'm saying we can look at their process and what they did. I think they talked to Jim West and Sandy O'Neill and I was talking to Fu Shin Lee a little bit on this. And I think we can just look at what's in EIM and various data sources. Because I think a just a threshold criterion is protection of human health and the environment. So I guess we clearly have to document it. That's my basic point. And we really have been looking largely at risk from bioaccumulative contaminants. But I also think we may be drafting issues of some sort to look at direct contact risk and ensuring that we are protective there. So we need to do some front end work before we continue on.

TERESA MICHELSEN The nice thing about all of this is that we have done most of this front end work, the RSET program. We already know which chemicals are detected in tissues. We did that whole pulling out of the data and comparison. We know what chemicals are connected and in what frequencies and sediments each of the Regions has chosen the bioaccumulative

contaminants of concern. We've got target tissue levels for protection of fish, human health and wildlife.

CHANCE ASHER Is this for fresh water though? Not Puget Sound.

TERESA MICHELSEN It doesn't matter. We did note fresh water versus marine and it made no statistical difference for bioaccumulative compounds. We have a whole set of criteria for how to define bioaccumulative contaminants and which ones that have been through the DMMP approval process. It would be a really quick exercise. In RSET, we didn't have the BOLD data yet so we couldn't do a comparison to background but we did do a comparison to detection limits. So we do have that part done and now that we have the BOLD data it will be easy to take the risk based levels and compare them.

LAURA INOUE That assumes that you can get from the sediments to the fish. At the tissue levels we defaulted on, were not touching the BSAF because that gets extremely site specific and troublesome. So we've got target tissue levels. And I did present that in one of the earlier meetings where these are the ones that even the target tissue levels are below detection limits and that contact could become more of an issue.

LON KISSINGER So that was the value of old Ecology work and old work done by PTI. They developed various regression relationship upper confidence limits on those regression relationships to develop screening sediment levels. You could use those BSAF to compare those levels.

LAURA INOUE I think ODEQ did that, and most of their values came up extremely low.

TERESA MICHELSEN It's important to note that if it's below detection in tissue levels, it's probably going to be below detection limits in sediments. And, every carcinogen, essentially, came out blue in detection limits, sometimes by orders of magnitude, but non-carcinogens, not always. You could potentially have a situation with a non-carcinogen where you were above background. But I think that for carcinogens, you can assume you're going to be below background.

DONNA PODGER We have done some of this work, and there's a lot of stuff we've done on these kinds of detailed type technical issues that we haven't presented to the group. We've been collecting this background information. It's not like we've been ignoring them. For us, we've run out of time to get into all of the details. I think that one of the biggest problems was the sediment/tissue relationship. It's fairly easy to calculate the tissue levels, but what does that mean for mercury, what does that mean for all these types of bioaccumulatives. You have to go and look at each individual compound, look at scientific literature, and there's orders of magnitude differences in BSAFs, especially for these summed compounds. It's not easy.

CLAY PATMONT No. On this concept, it's like we're making the assumption, we're going right to the sediment background, and that's probably where most of these sites will go. But, it might be worth preserving in the rules an option that if you have site specific tissue data and you can compare it with a reference. I'm thinking that mercury is a really good example of our ability to predict the methylation rates. And, there are other sites where, because of site specific attenuation factors, whether it's the black carbon or something else like that, it's not getting from the sediments to the fish.

CHANCE ASHER What if it's arsenic? You were talking before about the issue of Bellingham Bay. That correlation of sediment and tissue concentrations?

CLAY PATMONT I guess where I was going was – there's no downside to having a step in here that says – before you commit all these resources to go in and cleanup sediments, if you have site specific data that you can compare with a regional background or you choose some background. There may be some situations where you don't actually have any evidence of any bioaccumulation of level of concern.

DONNA PODGER I don't think we have any tissue background data.

CLAY PATMONT I know – there are some details there. But, just conceptually I'm just thinking, like in the Bellingham Bay area.

LAURA INOUE We have thought of it. And, the biggest problem is getting the reference tissue data together, because we really don't have that right now.

CLAY PATMONT But, there is the data.

TERESA MICHELSEN There's tons of data.

LON KISSINGER Actually, for the Puget Sound Naval Shipyard you're looking at some of the information.

DONNA PODGER If you could point us to where there's reference area tissue data.

PETE ADOLPHSON I think the point is the concept, just to allow the concept to remain. So that we can do that if we don't have the data now, but we can gather it later. Is that what I'm hearing?

CLAY PATMONT That's what I'm thinking. And, it's all going to be site specific. But, there are situations, and Bellingham Bay is a good one, where there is enough tissue data to say: do we really have accumulated dioxin in the tissue? There may be localized sediment concerns, but you go out and you look at the tissue, and it's not at all clear that there's actually an elevation going on.

TERESA MICHELSEN What about our traditional bioaccumulation test? I mean we always had those as overrides for DMMP. One could take a reference sample and run a bioaccumulation test and a site sample bioaccumulation test but we don't have the details.

LAURA INOUE Well no, this is what the DMMP is working on. It sounds reasonable but we haven't vetted it. We've acknowledged that in the recent guidelines that were just released that yes, we can allow a bioaccumulation test but we don't have the details of exactly how that works, including what is really different. How to tell statistical difference from the reference? That was part of the next series of work groups, not what the DMMP was talking about doing.

TERESA MICHELSEN DMMP has been doing bioaccumulation testing for awhile. It's not like it hasn't been done before.

LAURA INOUE We had target tissue levels. When you have target tissue levels that were below detection levels, what do you do?

TERESA MICHELSEN You use detection levels.

LON KISSINGER That was the main thing, documenting the fact that risk based levels were below background, so we could set the background as a regulatory criterion.

CHANCE ASHER Alright, any other comments right now on the document. Does anyone else need some more time to send written comments in? Everybody – I think – in this group was cc'd on all these comments. Everybody's got hard copies, and we'll have these electronically as well, so as you send more in, we'll make sure to make those available. We have another meeting scheduled for June 2<sup>nd</sup>. Most everybody's agreed that they can attend but some can't. Lon can't. But, I'd like some feedback from you about what we'd like to do at that meeting. We've got other issues that we can discuss. I think that we'll have at least some very rough language in terms of a narrative for the ecological risks from bioaccumulatives issue. We could continue discussions on the freshwater criteria if you'd like to do that. I'm not quite sure where we'll be at that point.

CLAY PATMONT When is the joint meeting – when are we doing the two groups together?

CHANCE ASHER The joint meeting is the 21<sup>st</sup> of June.

CLAY PATMONT So this is the last meeting before.

CHANCE ASHER Yes, and it would be good to hear from you. Before, we've always been setting the schedules and setting the agenda. Obviously, the human health background issue is big, but I'd like some guidance from you guys on what you'd like us to focus on for that meeting. And, it doesn't necessarily need to be today to come up with ideas, because we need time to prepare as well.

GLEN ST AMANT I guess we started this process by having some ideas and options proposed by Ecology and we talked about a number of different issues. I guess where I was hoping this would go is to get some sort of indication of Ecology's decisions by them revising at least some of the concepts based on the discussions the groups had. That's where I saw this eventually saw this process going. So, I appreciate the effort and the interest of characterizing what the group has said.

CHANCE ASHER Let you know what we are thinking?

GLEN ST AMANT Yes, that's what I'm hoping for.

CHANCE ASHER We'll try to go there as best we can. We'll definitely keep that in mind. I think we may need some more time on the human health background issue. Perhaps we can work on a few of these issues for the human health background issue for June 2<sup>nd</sup>. Perhaps the definition of Regional Background and if there are some other hot topic ones, rather than the whole thing.

CLAY PATMONT We probably need to hear from you a little bit, at some point anyway. Are there things that you would like to bounce off us? Are there some things with specifics around the regional background definition or something like that? I think we're all willing to help you in whatever way we can, within time limitations.

CHANCE ASHER OK. That's good to hear. This document was a start. And having this discussion.

LON KISSINGER We were talking about some site specific examples – just picking some of the urban bays and looking at, say, some of the major contaminants, PCBs, dioxins, furans, mercury – that might be useful, too.

LAURA INOUE Didn't I try to do that already?

CHANCE ASHER What I'm hearing is to have more specific questions that we ask with that data.

LAURA INOUE Harbor Island was one example, but Elliott Bay was another example, and those are the best characterized areas we could get a hold of. It's efficient to do things for PCBs and Aroclors because by the time you try to explain the sources, you had less than two dozen data points. All data collected is based on defining a site, so you go back to Greg Glass saying that you can't use existing data, because the data was collected for a completely different purpose. In order to develop a regional background, you have to sample for those sites that you consider in the regional background instead of using data that was sampled trying to define a site.

DONNA PODGER PSAMP went out and sampled the same location year after year.

CLAY PATMONT By the way – a different example, because Port Gardner had a very rich data set, and a number of other bays out there actually have more data than the two you just mentioned.

CHANCE ASHER Yes, I think I'm hearing a little bit different question, because we have done 6 bay-wide studies so maybe we can look at them.

LAURA INOUE It's my understanding that you maybe didn't have full chemistry for a lot of those sites, though. That was one of my original suggestions, using one of those baywide studies, because it had so many data points, but I was told that there was a tiered chemistry on that and not all the samples were amortized.

CLAY PATMONT It's not perfect, but having looked at a few of those, it's more than Elliott Bay. I don't know how much more – a factor of 2 or 4.

CHANCE ASHER What questions would we be asking and trying to answer if we present some of the bay-wide data. Say we looked at Port Gardner.

CLAY PATMONT Try building on what Pete was saying a consistent view of regional background, how would it play out? That might be one question.

GLEN ST AMANT So am I the only one who gets a little confused about whatever regional background is? Let's say you have defined that and if you do find that you move ahead and you've identified a site you want to clean up on a new plot. That's the regional background number to that site and an area weighted average - how do you actually apply it to that site as that next layer down in the onion? Maybe everybody else has a good idea of how that is approached but it seems significant.

CLAY PATMONT We skipped over that. It was going maybe a little too detailed in that how it would work in terms of defining a home range. A reasonable low end of a home range of the species of interest then looking at surface weighted average concentrations across that. Obviously the details would need to be worked out in reasonable exposure for the species. That's the way I see it. And try to map out a little bit of a process, again with the caveat and there are a lot of details that are very important. It's more the concepts that are the key thing right now rather than the mechanism of how it works.

LON KISSINGER Another approach might be to develop kind of the fingerprint boundary for the site and take out enough of the contamination so that the 95% of the upper confidence limit on the SWAC falls below the 90th percentile.

TERESA MICHELSEN Yes, home range introduces enough problems that we need an alternative approach. If you want Elliot Bay to be on an area weighted basis, every organism has a home range, how can you pick one? But if you want all of Elliott Bay to meet that level ultimately then each cleanup site has to meet the level for its own area. And that would be a more simple way of saying to a PLP what they have to do. That could be one option. But there's different ways. I would put out another alternative for June 2. I think that by June 21, we'd love to hear back from Ecology where you're going. We could all hear it at once and will talk about it. Maybe next time address some of the topics that we haven't really touched, like fish and other deleterious substances and statistics, and things that we haven't spent as much time on that were on our original list of topics. That's a possibility.

CHANCE ASHER Anyone else have strong opinions?

CLAY PATMONT The other deleterious substances, that kind of fell off.

CHANCE ASHER No, it didn't fall off intentionally. We are still working through it internally.

CLAY PATMONT That's what I heard. It just wasn't ready for prime time.

CHANCE ASHER Yes. It's become more complicated than we expected, and we have been working with the AG's office on rule language and we still have some internal discussion to do. It would seem like it's simple to do but it's not.

TERESA MICHELSEN Do you think you will be able to speak about it?

CHANCE ASHER Maybe, it's really hard for me to say right now. I am not intentionally trying to keep anything from you, it's just that we have a lot of internal discussion to do. So it's still a goal to keep it on the table. Any other comments about human health and background, or the June 2 meeting?

LON KISSINGER The methodology for 95% upper confidence levels on SWACs is not stable or well worked out. So that will be interesting.

CHANCE ASHER I'd like to open up to the audience to see if any of you folks have any questions or comments to make.

LAWRENCE MCCRONE I have a long list of comments. I'll just touch on a couple of them. Thank goodness we now have the BOLD data that we can use for natural background. Is that going to be good for Eastern Washington? Or Lake Roosevelt? Or the lower Columbia River? I don't know.

LAURA INOUE No.

LAWRENCE MCCRONE That's important. You've got to have reference data for different parts of the state. I think that issue came up about the practicability of defining regional

background. It really comes down to how many regions you expect to have. I can easily imagine three dozen separate regions around the state. Different reaches of the Columbia, Grays Harbor, Willapa Bay, Port Townsend, Port Angeles, all of the urban bays of Puget Sound, Lake Roosevelt, Snake River, Yakima River. And so forth. And you start thinking: do I need to define regional background for all of those? It becomes a much more difficult proposition. And one thing I would suggest as an alternative is the concept that maybe we just think in terms of setting something as a regional background for urban areas. It's not perfect but it's something that we don't need to define each region individually. I know that the examples stuff that you sent out was just examples but the difference between regional background values and natural background values that you have isn't nearly big enough to bother with. And if it's only a factor or two difference between regional background and natural background that doesn't matter. It's not worth fussing over ad infinitum about that.

One thing about the PLP's getting together to define regional background, I'd say that there are many PLP's that are going to be reluctant to define a regional background for, say, dioxins if they're not a source of dioxins. Or any other chemical you can think of. One thing about regional background is, a lesson that I've learned working on the Duwamish, is that regional background may not be characterized by going out and collecting a lot of samples today. Because what that represents is the sum total of a lot of sources, or a lot of cleanup sites that we're trying to do something about. And what we're coming to in the lower Duwamish for example, is trying to project what going to be the background in the future, not background based on current sampling. So that's a really sticky issue. And you can spend a lot of time thinking about that one.

I'll just mention one other. Reflecting back on the time of promulgation of the original set of sediment management standards, the completely separate economic was done at that time and my understanding is that will still be part of this rule revision process but it comes once you have a draft rule ready to go out for public review. I really think that some of the economic impact should be taken into consideration before that time because some of the things that we are talking about have such huge economic impacts that they will never ever fly.

CHANCE ASHER Is that related to any specific issue?

LAWRENCE MCCRONE I think in terms of work load for Ecology, some of the things we're talking about would represent an order magnitude increase in the sediment program for Ecology if you go statewide for every plan. Ecology is having struggles to fund what they're doing today, and I think what we're talking about is an enormous increase in the level of funding that would be required. But above and beyond that is the impact of the regulated community. I think the regulated community isn't going to be able to clean up the world to a natural background.

TODD THORNBURG Following up on some of the threads of discussion there, I just like to voice disagreement that I think there are aspects of source control that don't belong in MTCA or SMS. They really just don't fit into the framework. Just to throw out a couple of examples, and these are things that we are going to be having to deal with in the next couple of decades. It would be nice if we got some more hybrid vehicles on the road. If we do that we'll see a decrease in PAH levels in our road runoff. But this is a national issue. This is us as a nation changing our attitudes and our lifestyles. It really doesn't fit into the MTCA framework. Phthalates are evidently gassing off from vinyl building products. We are not going to go out and tear down everybody's vinyl siding and everybody's electrical conduit. Maybe at some point we as a nation

might want to come up with a product substitution for phthalates. But that involves a whole host of other decisions. What's the cost of that product? What's the manufacturing cost, the carbon footprint, and the host of other issues. To me all that kind of stuff is well beyond the framework of MTCA.

ROB PASTOROK I think that we recognized before by the Ecology panel that regional background might need to be redefined in the future. I'm not sure if you can project to the future what it's going to look like. I think basically existing data is ok. It's not all that easy. A point of clarification, Teresa, you said something to the effect that the bioaccumulation levels that were derived.

LON KISSINGER I was going to say, the consumption rates would certainly be different.

ROB PASTOROK Between freshwater or marine is that what I was hearing?

LONG KISSINGER No, you are talking about target tissue levels I think, weren't you? But consumption rates could be conceivably be different between marine versus freshwater environments although the Columbia River tribes consume at about the same rates.

ROB PASTOROK I was thinking about wildlife. They are entirely different receptors, so how could it be the same?

TERESA MICHELSEN OK so just a point of clarification then. RSET includes a variety of tables for like 15 different wildlife receptors. You are allowed to pick and choose according to your site. It also includes 3 separate levels for fish consumption for human health. So, we have target tissue levels for all those varying scenarios. And we, as RSET, because we were trying to serve three different states in a variety of different types of programs didn't try to dictate which organisms you should include and which consumption rates you should use. We just provided a variety of tables for different scenarios. So you can modify as you see fit; or Ecology can modify as they see fit.

ROB PASTOROK That's probably fine for now. It seems like there's a lot hinging on these risk based levels, especially for human health are below, or going to lead to below backgrounds. So we have got to really put all this emphasis on background. Where is Ecology on evaluation of risk assessment methods and approaches? Is that all set in stone or could that influence this driving force to work towards background, background, background, so much. You know it matters.

CHANCE ASHER So the question is, if I understand it, depending on how you approach the risk assessment, the numbers can come out different and what is Ecology doing about that right now?

ROB PASTOROK Right. Is there any parallel evaluation of risk assessment methods for developing bioaccumulation based values, for instance?

CHANCE ASHER Well not in terms of developing numbers criteria, we're not going there. But we are looking at trying to developing processes in guidance on how you do a risk assessment. We were dealing with fish consumption rates as well so hopefully all of that will coalesce as best

we can with promulgation of the rule. But we are looking at developing some type of guidance that provides some specificity on how that's done. The big question is how detailed should that be and dealing with the issues of workload.

LON KISSINGER I think even at like moderate seafood consumption rates, you're going to see things are below background, certainly for PCBs.

LAURA INOUE Most of them even for recreation of 64 grams was below detection limit for the nasty bioaccumulatives.

TERESA MICHELSEN Knowing these folks as I do, I think they were probably looking for more profound changes in the paradigm that's used. And the answer is probably "no" because of Ecology and EPA's existing methodologies. I mean that's at least that's the response I've had to give in the RSET process because I didn't have the option of changing those.

CHANCE ASHER We don't have a consistent process on how to do that for sediments so I can't say the answer is "no." I can say the goal is always to make, whatever changes we make or guidance that's put out for sediment, is as consistent with MTCA processes as possible and as applicable. And when we're looking at fish consumption rates we need to definitely deal with that consistently. Dave, want to comment?

DAVE BRADLEY I don't think these are profound, but to me there's a wide range of issues with the risk management and risk assessment framework. I think the big changes will come if we started to look at different risk levels, which right now that's not the part that we've opened up with the rules. I think with the narrative and human health SMS there are some issues there but you get into the consistency with the MTCA and WQ standards which can kind of constrains things. So the whole methodology of how you characterize what is a particular risk level, there's a lot more guidance issues to be looked at in the sediment. So it's more the kind of characterization and policy types of things.

Even things allowing for fractional site use which is done commonly in risk assessments. But that's the way it's being implemented in WA State, there's no allowance for that. You have to assume that no matter how small your site is that you could harvest all of your seafood from there time and time again, whether or not it's any realistic expectation that that site is sustainable.

DAVE BRADLEY And again I think that's, it's how we characterize. We make a lot of assumptions, we manage uncertainty in different ways and I think there are a lot of things coming together, and I would agree with Clay in a weaker moment. You need to look at what's the basis for the standard when you're looking at how you measure compliance. And I know some people are enthused about the surface weighted averaging but I think kind of looking at how that matches up, what's the basis for the standard, I think we need to be looking. So I guess there are some devils in the details with the characterization, but in sort of the broader risk management, we're not really looking at opening up those issues.

CHANCE ASHER I need to move the comments along. Heather?

HEATHER TRIM So in contrast to one of the comments made earlier today in this committee, I thought one of the best comments made in the advisory committee last week was the issue for

these urban bays. We do need to know the fate of the chemicals. It's not just a matter of just being further away from the sources or not it really depends on the bay. So I guess we would be advocating for chemistry characterization. For doing it right because it's an excellent idea. I think we should put together, at the end, a list of recommendations to the Puget Sound Partnership to include in their action agenda. They're going to be updating the action agenda either later this year or early next year and there are very weak recommendations in there right now with regard to sediments. And I think that's because there was such a big deal that everyone was working on that they just couldn't put so much effort. But there's a lot of detail about some topics but not this topic. And this issue about some of regional background would be appropriate. Maybe it could get federal funding?

An issue in terms of Elliott Bay which seems to have not been brought up yet, which makes Elliott Bay even more difficult to do this assessment for, is that we have the two largest sewage treatment plants, 50% of the treated sewage coming in at each edge of it. So we have discharges at Alki and at Discovery Point. And that's from all of King County, so that's not representing some urban signature right there. This is representing all of King County with a huge amount of industrial waste. So in terms of Elliott Bay and this regional background, there is a pretty major complication right there that needs to be taken into account.

CHANCE ASHER Can I clarify? Are you saying that that type of source should not be included in a regional background calculation?

HEATHER TRIM Well, if you're trying to get away from sources, you have two huge sources out at the ends so if you're talking about the central basin.

Yes. So it really does complicate it unfortunately. And of course these chemicals are the chemicals that would be there. The unfortunate thing about the Clean Water Act is that it does not have these kinds of triggers for storm water. So right now the amendment that was done in, what 85, for storm water does not address source control in the way you have it addressed for municipal discharges. So it does need to follow the MTCA to deal with this. As Teresa pointed out, there is language related to this so we can't rely on the CWA to do this. That's just not in the current language. And then the last thing is, on this issue of the recommendation for site specific, it just gets us back to where we were before where we have fights and all that. I like the fact that Ecology is trying hard to come up with something that's consistent so we don't have a fight at every site. So I have to say I don't think that's a great idea to go and have an exception. If there's an exception, it's got to be very carefully crafted, or we'll get just get back to the same.

CHANCE ASHER Thank you. Anyone else? I'd like us to move on to the last piece of business which is a document you have that's at the end your handout. I'm going to let Pete take over.

PETE ADOLPHSON We are trying to integrate the terms that are in MTCA and SMS to essentially clarify and make it simpler when the people are looking at both rules. Obviously, it's not terribly easy to do and they're not going to mesh perfectly well. So the goal was just to try to bring them together as closely as possible.

The first one deals with RI/FS integration on how alternative approaches are met in both MTCA and SMS. And we essentially kept everything that was in SMS, in terms of the evaluation of

alternatives and then added one term not in SMS that came from MTCA. That was the permanence aspect. I simply pulled the permanence language that was from MTCA and pulled it over into the SMS language. I referenced the MTCA section. You can look at all the other ones that are currently in SMS and see how that flows but I don't want to go through these in too much detail. But you can think on that and then go back to both rules and see how that flows.

The next section is the terms and I'm not going to go through all those because there are several pages of terms. Some of them are new, some are just redefined terms and this again is for clarification purposes. Some of the terms have examples associated with them to try to clarify what those terms are. There seems to be some confusion in some of the old terms and how they were used. We had some conversations earlier today about the term "chronic." I didn't include "sub lethal" but I tried to include some terms of what chronic and acute actually are. And you can see how, when Russ had the discussion about freshwater, chronic and acute, how I redefined them. I want to make sure you guys understand these are just proposed. And so if you have comments about what you would include or exclude to clarify the rule, I want those comments and we'd be happy to incorporate any of those comments as well.

Interestingly, when the SMS rule was first put together, the term "sediment" was never defined. I found that interesting. So Russ and I split apart those two terms because they're not the same thing. We put those together. What's interesting is we have locations where, especially in freshwater where you'll have an ephemeral pond or some sort of freshwater marsh, and half the year it's dry, and we have a lot of questions, well is that sediment or not? And so we tried to incorporate that kind of idea into the definition of sediment. Not all the comments, even the internal comments, are included in this yet. So this was just thrown out so that people could make comments.

CLAY PATMONT What's the general basis for six weeks?

PETE ADOLPHSON Russ and I had this conversation too. He looked at the definition of a wetland, and that didn't help. My thought was we were looking at the effectiveness in most systems on invertebrate communities. And I went to adult insects laying eggs that are going through their development and that time span. So it was essentially a six-week period. So that's where that came from.

CLAY PATMONT That would include things like X for example, because it would deposit, hang out for awhile before it moves ... it would include things like that.

PETE ADOLPHSON That is correct. The way I try to go through these, is by example and think of an example and does it fit or not. And I'm sure you guys can come up with a lot more examples than I did and that's why I put it out there for you to sort of cogitate.

TERESA MICHELSEN What is the use of the term "active cleanup action?"

PETE ADOLPHSON It is used elsewhere in the rule, active cleanup action is actually used in the rule.

TERESA MICHELSEN Where?

PETE ADOLPHSON I think it's under either minimum cleanup level definitions. I think it's one of the last three where I said we're still working through the process.

TERESA MICHELSEN So it's kind of like you have to do an active cleanup action to get to level X or whatever.

PETE ADOLPHSON Yes, exactly.

TERESA MICHELSEN I just wondered what it was for.

PETE ADOLPHSON And probably the best way to go through some of these is if you have an electronic version of the rule, is to search through these terms and to see if they fit in each case. Which hasn't been done fully yet even internally. But those are the types of things I want you guys to look through. And we need to do the same thing. I tried to do it in most cases but it's hard to do because it refers back to other sections of the rule you know. It's not a linear process so that's why it's difficult and that's why I need your help to go through it as well. The, also "biologically active zone" was never defined in the rule. I think that's pretty important. And just examples are included in this definition as well.

TERESA MICHELSEN I would be more specific about a "majority of". I think in the original study it was like 90% or 95%.

PETE ADOLPHSON Yes, we can provide more detail, if you can provide written comments if you know if there should be some sort of number associated with it, so it's less ambiguous, that would be good as well.

TERESA MICHELSEN Those things I would say by default is the sediments in marine environments, there's no data for freshwater and it's random as you know. So, just to clarify that.

PETE ADOLPHSON Good point.

ROB PASTOROK And just be clear whether you're talking numbers or biomass.

PETE ADOLPHSON In terms of what biomass?

ROB PASTOROK Just be clear about whether majority means in terms of numbers or related to biomass.

PETE ADOLPHSON Right. OK. Good comment too.

CLAY PATMONT On the active cleanup action, if you're defining sediments the way you define them now, as six contiguous weeks, it changes, at least in my mind. There are a lot of sites that the active part, the expensive part, is the source control on a facility which can be very, very complicated. I think historically it's a much bigger issue than has been on source control. I'm just thinking you'd want to write this so you didn't box yourselves in. So that you've cleaned up all your sources and you've got 6 week's worth of sediment, that's probably an extreme example. But to me active sediment is and active cleanup, might be active source

control. And that might, you know, do you feel the rule constrains you against that? There are sites where sediments responded very quickly to a source control and it doesn't make any sense to go in with, if it's sort of a dredging and capping.

CHANCE ASHER Let's go into detail by comments either email or a phone call to Pete, so Pete can capture that. He has to document any conversations that he's got.

The goal is to give you guys some work outside of the meeting today. We actually need to be out of here, everybody has to be out of the gate by 5:00. The goal is to try to harmonize the rules and so that when you you comply with the SMS you comply with the MTCA law and you don't have to toggle back and forth between the rules to know how to comply. So that's the really the major goal. Then we found that there's fuzziness in the SMS rule itself and conflicts with some pieces that we've found in MTCA. So that's the overall or arching reason for this piece of homework for you.

We'll have it online for you. And so I'd like to give people some due dates and give reminders of what we've all talked about in terms of homework. So we have comments on the freshwater SQV report as well as any comments you have about the biological criteria framework. I'd like to have that by May 17. We need to be able to incorporate that. We're going to have a discussion with the larger MTCA/SMS group as well. I think for the Integration piece that would be great if we could get comments by that date as well. Because we really do need to turn this around pretty quickly because twice a month we've got these major meetings. It's not just you all. We need to turn work around and let them know what kind of the things we've gotten from you and where we're going. And then there's the summary document and some folks still want to provide comments on that. So if we can get all of that by the 17<sup>th</sup> of May. Two weeks from today.

LAURA INOUE Also we only have four days before the next meeting, so it only gives us a day or two to collate everything and get it back out to everyone.

CHANCE ASHER So if you could get that to us sooner, that would be really great. But I understand, we all actually have lives and work that you're getting paid to do. And this is all essentially free, pro bono. So let me take a look and make sure I caught everything. Then Russ and Teresa are going to synthesize some of the questions that we had, that we were bringing up on the biological piece for the freshwater criteria. They'll send to the group that you get to answer as well by the 17<sup>th</sup>. So, you get 4 homework items. Teresa gets 5. Sorry. Any last questions, comments? Well thank you very much for your time. We appreciate it. We kept you 20 minutes over the usual time and I appreciate it very much even though I scheduled it until 4:30 and no one argued with me when I sent it out so that was great. Thanks again.

## **Sediment Workgroup Flipchart Notes**

### **May 3, 2010**

TERESA MICHELSEN – question about calling *Chironomus* acute/chronic – only 30 data points. Criteria not developed yet for *Chironomus*. Freshwater species difficult to get appropriate conditions. May be difficult to get three freshwater species.

Option 1 – keep the acute and chronic tests but have three distinct endpoints with two or more species. Chronic tests can have two endpoints.

Pete Adolphson - if one test with two endpoints? Establish criteria for several endpoints.

TERESA MICHELSEN – use both for calculating guidelines

RUSS MCMILLAN – SQV report – top five used. With freshwater, we are trying to mimic marine water tests but may not be able to. We are trying to establish one chronic and two acute tests for freshwater.

TERESA MICHELSEN – sub lethal endpoints versus chronic tests. May include 10 day growth, but this really isn't a chronic test

PETE ADOLPHSON – look at life cycle of organism; defining acute and chronic

TERESA MICHELSEN – keep the terms straight for regulations

CLAY PATMONT – 10 day versus 28 day *Hyaella*. This is a national dialogue with EPA. Some believe that 10 days is not long enough. A lot of discussion on this topic.

LON KISSINGER – want to take this information back to Burt Shepherd and get his take.

CLAY PATMONT – how references compare

RUSS MCMILLAN – sample area – come up with reasonable values.

CLAY PATMONT – reference pool highly uncertain; ways to clarify; don't know if can meet them. Tox – explain but maybe an artifact.

RUSS MCMILLAN – challenge – Reference QA/QC is required. Are we looking at true toxicity?

TERESA MICHELSEN – reference discussions moving away from t tests.

CLAY PATMONT – huge issue; need high level of understanding.

TERESA MICHELSEN – if criteria promulgate.

CHANCE ASHER – adjust false negative rates; not choosing the most conservative rates – policy issue. Trying to stay similar to framework like marine version confirmatory tests.

RUSS MCMILLAN – quite a challenge. No toxicity versus lower range and false positives; balance to the 20% as close as possible; finding best balance of false positives.

TERESA MICHELSEN – False positives have greater number of stations than false negative. Never make decision on only one station. Minimum of 3 stations, so 5% chance to fail to design a site that should be other fail safes. Dredging project is a little different.

PETE ADOLPHSON – that’s an important point – certain freshwater sites with unique situations; look at buffering, seasonality, etc.; still following marine paradigm; important to have balance.

TERESA MICHELSEN – critical for freshwater more than marine; concerned about validation tests; eventually we will / increased variability at these sites (freshwater).

LAURA INOUE – frustrated with data set; couldn’t get QA2 level.

TERESA MICHELSEN – only a few stations; a couple of labs – weren’t able to get that information. worth gathering; variability issues – chemicals, etc. be reliable on programmatic basis; be more flexible than marine.

RUSS MCMILLAN – challenge to/use/write ranges of results; capture as best as we can; may not be applicable throughout the state.

TERESA MICHELSEN – put in a couple of examples, unusual matrices/sites.

GLEN ST AMANT – other eco-receptors? Site by site?

CHA – fold in with bioaccumulation. Bioaccumulation piece doesn’t really impact these numbers.

LAURA INOUE – Working with RSET; assume these are benthic organisms; moving toward a number protective of juvenile salmonids; unable to get decision from NOAA; DDT further behind; not far along to develop standards.

GLEN ST AMANT – what will it look like?

CHANCE ASHER – more on June 2. Stay consistent to current framework; develop from regional data not lab/real world data. Concerns with east vs. west fresh water difference? Not sure, may be types of contaminants.

TERESA MICHELSEN – did test east vs. west; combined was greater reliability. Why? No statistical reason to separate data.

CLAY PATMONT – Dry weight basis? TPH – really parameter? What’s best measure of oil?

LAURA INOUE – yes, I think so.

CLAY PATMONT – huge advantage in last 5 years; a lot of good information; bioavailability – right parameters.

TERESA MICHELSEN – organic carbon – looked at LA Harbor normalized marine AETS; organic carbon normalized; no real difference with reliability. During rule implementation – probably with organic carbons. Some sites more or less organic carbon. Hard for some to understand; implementation issues. Look at several times – no appreciable differences. Sticking with dry weight.

LAURA INOUE – At national SETAC meeting – several talks; at least 5 different types of organic carbons.

CLAY PATMONT – Organic carbon normalized; useful for screen; have you thought about next step?

LAURA INOUE – Pore water; look at better for metals; real availability of these contaminants? Surfactants, etc. – assists desorption

CLAY PATMONT – trying to come up with a better endpoint; thinking practical; what is the next step?

RUSS MCMILLAN – bring stuff to Ecology

TERESA MICHELSEN – historically, site managers are not receptive – only works for certain metals; doesn't always work. Many variables/destabilized; haven't used much at NW sites.

PETE ADOLPHSON – fresh water systems only? At Chesapeake Bay - bounced everywhere; theory sounds good but couldn't find trend.

CLAY PATMONT – Dave Mount – has never seen fresh water site with metals toxicity; mostly at mining sites; these are tools.

TERESA MICHELSEN – when model gives metal values, higher than anticipated. TEL values are in the minimum and fresh water results are different. Looked at many methods – TPH – higher reliable results over all other analyses. TPH is key analysis.

TERESA MICHELSEN – cleanup methods so get contaminants and not decomposing leaves

TODD THORNBURG – RSET continued to look at – compile all info they could; looking at TPH, (liter) EPH (for EPH, virtually no data in sediment sites); heavier fraction VPH. Develop guideline for TPH and try to correlate EPH – no correlation found (lousy coefficients). EPH – fewer labs and analysis issues; but more information than TPH. TPH is more reliable.

TERESA MICHELSEN – Simpler approaches work just as well as more complex theories; easier to implement.

ROB PASTOROK - noise maybe due to cross-section of data; option of increased complex approaches could be valuable; allow site-specific data. Bioassays – sometimes try to...

PETE ADOLPHSON – approach – heterogeneity; 40 reference/site samples/seasonal/5 years.

ROB PASTOROK – special studies; how do they relate?

TODD THORNBURG– good work, Teresa; big step forward. Big challenges = bis-phthalate

LAWRENCE MCCRONE - Heard Ecology staff may make decisions based on one sample; no, must have three samples; make decision to list = cleanup; would require at least three stations.

LAURA INOUE – site – decide within the site?

CHANCE ASHER – hesitate to comment; this is site specific.

ROB PASTOROK – multiple solutions; does it matter in what order they are entered? Brings false positives down; refined algorithm; look at co-variance; then take lower value and work up.

TERESA MICHELSEN – start low and work up to address co-variance – works well, but not perfect.

COMMENTS:

JOANNE SNARSKI – haven't sent comments yet; good attempt – wasn't shocking; captured the information; surprised by the amount of comments; some didn't understand.

GLEN ST AMANT – weave comments in.

PETE RUDE – technical challenge – how to get through it. How to get it all done in a reasonable time; struggle with regional background; numbers may not help move forward.

CLAY PATMONT – regional background; could be huge task / agreement – be more specific; regional background seems to hold promise. Important for definition; “hydrodynamic” sounds good until we try to define it. More watershed scale; huge complexity; regional background cannot be influenced by facility. Upland site as defined in MTCA.

PETE ADOLPHSON – need to understand – influence.

LON KISSINGER – diffuse PCBs in Elliot Bay.

TERESA MICHELSEN – two points:

1. How do we gather all this data? Interagency coordination. We need urban bay action teams and data collection; need funds to do this; data collection; can carry out this.
2. Don't be afraid to use professional judgment; define at drop off?

Years ago, data/GIS data were looked at and discussed – rationalize.

TERESA MICHELSEN – right now we’re stuck. Need regional background data; rule could incorporate that; could have cases for various; that is, site by site.

PETE ADOLPHSON – We would need to define regional background before collection?

TERESA MICHELSEN – yes, write up before as did with BOLD data; made steps and followed them.

CLAY PATMONT – how do you incentivize willing PLPs? May not agree with theory but probably with geographical boundary; many ways to address.

CLAY PATMONT – reasonable PLP that wants to do cleanup; cooperative willing parties (versus recalculation); use small amount of data?

TERESA MICHELSEN – Pete Adolphson’s comment “What’s state of Chesapeake Bay?” Take out of polarized context.

GLEN ST AMANT – agrees.

LON KISSINGER – want another statistical discussion.

CLAY PATMONT – First identify natural background – define number; go to embayment and where is that number? If concentrations are down, then look at BOLD data; how to deal with?

TERESA MICHELSEN – applies to areas where data collected.

LAURA INOUYE – gathering data; Straits of Juan de Fuca – no disposal sites there; BOLD more on/around urban areas; only rock further out.

CLAY PATMONT – BOLD data worthwhile – 90% - a good place to start. How to determine within the influence of the site? Value judgment; keep simple/simple tools; start conservatively.

LON KISSINGER – why pull out areas of low natural background?

TERESA MICHELSEN – outer boundary – a lot of Commencement Bay meets natural background. To draw boundaries – natural vs. regional background. Don’t apply regulations to cleaner natural background areas. Do we include those points within regional background data? Looking for break point vs. boundary – map data – how to define sites more; see if it’s complicated?

CLAY PATMONT –

1. Elliot Bay most complicated; plus some others
2. Slightly less complicated than others
3. Then look at simpler bays

Don’t want to drag them through Elliot Bay type scenario.

CLAY PATMONT – Best Professional Judgment (BPJ) – regional background – “preliminary cleanup level” that is, Elliot Bay – all PLPs probably won’t all work together; depends on complexity of site.

TERESA MICHELSEN – don’t forget about the tools we already have; just because Elliott Bay (for example 10 discrete sites and not all just one site); site unit = engineering term; all separate sites. Stick with all separate sites – keep same structure.

PETE ADOLPHSON – separate sites / mixed plume.

TERESA MICHELSEN – watch the terminology. Site list – define as separate site. How it’s handled – each PLP resolve; liable – final buyout and mixed plume is monitored.

TERESA MICHELSEN – calculate net present value of monitoring cost.

CHANCE ASHER – same concept; same issues there. How to get cleaned up vs. mixed plume; above natural background. Larger site / mixed plume; concepts are same.

TERESA MICHELSEN – distinguish between sites – imp.

CHANCE ASHER – what we’ve heard.

CLAY PATMONT – identify units that make sense for the medium; PLP – Ecology should come to us with what makes sense. Recontamination is complicated even with samples – more complicated BPJ.

CLAY PATMONT – hot/warm spots – above natural background; might not agree with parts, but agree action is needed now; for example, limit recontamination.

LON KISSINGER – trust funds; sequencing of remediated sites; estimated for sufficient funds? Issue for DW. Does it make sense – worst contaminant to least?

TERESA MICHELSEN – societal issues; which – where to put money? Legacy problem. PLP shouldn’t worry about City of Seattle outfalls. Not originally PLPs issue. Dealt with legacy issues separately.

LON KISSINGER – changing public perception.

JOANNE SNARSKI – complex issue; indemnify Cascade Pole; work with them to come together; look at other sources.

TERESA MICHELSEN – ports are typically not responsible; such as Cascade Pole is responsible, not Port of Olympia.

JOANNE SNARSKI – definitely interesting – how to move forward and take phased approach; concepts in framework; short and long-term goals; different from small landowners. How to get phased in – details?

CLAY PATMONT – role of CWA and other programs; MTCA not designed for source control. Ongoing stormwater sources from large areas are the issue.

CLAY PATMONT – source control, stormwater, natural background – outside feasibility of MTCA/SMS. Area huge issues; societal issue.

TERESA MICHELSEN – made same comments

- Cleanup – legacy areas
- Source control – decrease large area.

Source controls aren't used the way they used to be; needs to be reinvigorated. Don't take to natural background within MTCA; not what MTCA is designed to do. Use MTCA for legacy sites.

GLEN ST AMANT – next phase – CWA? SMS? ongoing discharges? Stormwater?

TERESA MICHELSEN – either through MTCA/CWA

CLAY PATMONT – covenant not to sue; site-specific determination; other stuff out of control.

TERESA MICHELSEN – other regulatory authorities; be very careful with the language used.

LON KISSINGER – talk to people working on TMDLs; been used around Region 10 but not Washington. TMDL – all sources to the body of water are calculated; loads/contaminants; may be helpful to talk.

TERESA MICHELSEN – recently TMDL – taking into account sediments; could be used in Region 10.

LON KISSINGER – need to look at and compare risk-based concentrations and background - tissue? sediments? Decide which contaminants have limited distribution and higher background. Direct risk-based...

LON KISSINGER – justify it; seafood consumption advisories – did some of this analyze – Look at DOH's process. What's in EIM? Risks from bioaccumulation – but also look at direct contact risk.

TERESA MICHELSEN – a lot of this was done at RSET; bioaccumulation – no difference between marine vs. fresh water; RSET – compare to detection limits.

LAURA INOUE – RSET didn't touch BSAF

LON KISSINGER – PTI developed BSAF levels; compare.

TERESA MICHELSEN – If lower detection in tissue, then lower detection in sediments.

TERESA MICHELSEN – carcinogens are usually decreased in background.

DONNA PODGER – have done some of this that hasn't been presented yet. Sediment – tissue relationships – what does that mean for Hg? PAH? Look at each individual compound.

CLAY PATMONT – seem to assume – background; preserve in rule – e.g. Hg; before resources, if site specific tissue data to compare with background.

DONNA PODGER – Need tissue – background data.

LAURA INOUYE – reference tissue data?

TERESA MICHELSEN – look at fresh water data.

CLAY PATMONT – do we have accumulated contamination in tissue?

TERESA MICHELSEN – traditional bioaccumulation test? What about bioaccumulation?

LAURA INOUYE – don't have the details; DMMP.

LON KISSINGER – Risk lower background and skip to background.

CHANCE ASHER - What to cover at the Joint meeting on June 21

GLEN ST AMANT – is Ecology revising based on concepts heard here? How/what thinking?

CLAY PATMONT – we need to hear from Ecology.

LON KISSINGER – site-specific examples would be useful.

CLAY PATMONT – try different examples; how would this play out?

GLEN ST AMANT – regional background – site to move forward; how to apply regional background to that site.

CLAY PATMONT – define a “home range.”

CLAY PATMONT – tried to map out a process; concepts are key right now.

LON KISSINGER – fingerprint boundaries for site; not stable/now well worked out. SWAC – 98<sup>th</sup> percentile.

TERESA MICHELSEN – Elliot bay area weighted average; each cleanup site needs to meet that level; 1 option.

Another alternative – hear back from Ecology about where going; fish; other deleterious substances; other topics.

**Audience comments** (May 3 meeting)

EXPONENT– can the BOLD data be used for Eastern Washington? Natural background for different parts of state; how many regions around the state?

Suggest set “something” as regional background for urban areas but don’t define for each urban area. PLPs – regional background – won’t want to define this. Regional background – may not be defined by “today’s samples”; project – what will future background be? Original SMS – economic impact assessment – after draft rule; economic impacts take into account now; even workload for Ecology; talking enormous increase in impact to PLP/Ecology.

TODD THORNBURG – source control – some parts not for MTCA/SMS; need to deal with this. National issue – doesn’t fit MTCA framework.

ROB PASTOROK – regional background may have to be redefined in the future; not easy.

TERESA MICHELSEN– RSET exercise; bioaccumulation – target tissue levels; wildlife receptors – can pick; target tissue levels; which organisms/consumption rates.

ROB PASTOROK – Ecology on risk assessment methods? Could it influence – looking at background; how to approach the risk assessment methods to develop bioaccumulation numbers.

LON KISSINGER – modify seafood consumption rates.

LAURA INOUE –much lower detection limits.

CHANCE ASHER – no consistent process

DAVE BRADLEY – wide range of issues; consistency issues; methodology; guidance. what’s basis for standard? How matches with basis for standards?

HEATHER TRIM –

- do need to know fate and transport of chemicals; put together recommendation to Puget Sound Partnership - especially regional background; federal funding.
- Elliot Bay – 2 largest wastewater treatment plants – Alki and all King County outfalls to Elliot Bay; sources to Elliot Bay.
- CWA – doesn’t have stormwater triggers – does fall to MTCA; cannot rely on CWA.
- No exceptions or back to same.