UNITED STATES DEPARTMENT OF COMMERCE
National Oceanic and Atmospheric Administration
National Marine Fisheries Service
P.O. Box 21668
Juneau, Alaska 99802-1668

December 14, 2004

Ms. Nancy Sonafrank, Section Manager
Alaska Department of Environmental Conservation
Water Quality Assessment and Monitoring Program
610 University Avenue
Fairbanks, Alaska 99709

Dear Ms. Sonafrank:

In response to the Alaska Department of Environmental Conservation's request the National Marine Fisheries Service has reviewed the draft document reviewing literature on the acute and chronic toxicity of hydrocarbons specific to Alaska species. Enclosed please find comments specific to the draft document prepared by scientists from our Auke Bay Laboratory.

Should you have any questions regarding these comments please contact Ms. Jeanne Hanson at (907) 271-3029.

Sincerely,

For James W. Balsiger
Administrator, Alaska Region

Enclosure

cc: NMFS FAKR/HCD – Susan Walker
NMFS Auke Bay Lab – Jeep Rice
Records – Lori Durall

USFWS, ADNR-OHMP, ADEC – Anchorage

Lisa McGuire
U.S. EPA, Region 10
Water Quality Standards and Planning Unit
1200 Sixth Avenue
Seattle, Washington 98101

Comments on the Alaska Department of Environmental Conservation's "Acute and Chronic Toxicity of Hydrocarbons with an Emphasis on Alaska Species: A Review of the Literature" Alaska Department of Environmental Conservation

The National Marine Fisheries Service (NMFS) reviewed the draft summary of literature on the acute and chronic toxicity of hydrocarbons specific to Alaska species. The review of the literature relating to polynuclear aromatic hydrocarbons (PAH) prepared by ADEC in support of the triennial review of Alaska state water quality standards is an appropriate start for the triennial review. While it has appropriate objectives and strategies, it does not distinguish between several important concepts. The document is primarily an annotated bibliography, and more of a synthesis is needed prior to decisions being made to strengthen or weaken the current standards.

The document correctly divides information into "pre 1995" and "post 1995" papers. There has been a lot of new information relative to PAH in that time, including the Exxon oil spill, and an evolution away from a dependence on acute narcosis toxicity models. The significance of these major changes does not appear in the "synthesis" aspects of the document. Ultimately, the goal is to distill hundreds of studies, exposures, mechanisms, etc. down to an appropriate water quality standard that can be understood and lived with. However, prior to distilling the accumulated knowledge base to that standard, many important complexities must be considered.

NMFS review of the document encompasses general comments on the complexity of oil, analysis, units, and toxicity; evolution of acute versus toxicity mechanisms, and how they relate to water quality standards; strategy and accuracy in sorting out research into key tables; list of critical missing papers; and specific comments on several reviews of the scientific papers.

General Comments

PAH is a complex mixture

The review does not tie the chemical information given to the problems resulting from having a complex mixture such as PAH. Because there is a complex mixture of many different hydrocarbons in spilled oil, with many different PAH compounds within that mixture with attendant analytical complexity. Any analysis of spilled oil needs to consider a variety of factors including the amount of oil spilled, visible as well as the amount in the water column; chemical composition; various conditions, such as temperature, salinity, mixing energy and dispersant conditions; and what fractions and compounds are responsible for toxicity.

The pre 1995 literature has a mixed set of methods for exposure and analysis, and it is difficult to compare concentrations and compositions from one study to the next. Reporting concentrations as a "ppm" does not mean it is realistic or comparable, and largely depends on the methods of analysis and the composition of the PAH solution. The tables of data listed in the review contain a mixture of concentrations for different types of solutions [Table 1 for example contains data on PAH, water-accommodated fraction (WAF), pure compounds, with/without different

dispersants]. It is critical when comparing concentration numbers to understand what is being compared. Different analytical procedures may come up with similar units, such as ppm, but the composition may not be comparable. This is especially true when comparing TPH and TPAH (total petroleum hydrocarbons versus total polycyclic aromatic hydrocarbons). In Table 1, concentrations by Little are for TPH and there is no basis for comparing to the Neff citation where the concentrations are for TPAH.

Pre 1995 studies usually do not adequately quantify the toxicant solutions adequately. Different forms of mixing and analyses yield fundamentally different results. These concepts that can complicate interpretations across studies have not been treated in the document. Therefore, listing the combined results in the tables can be very misleading.

Evolution of environmental toxicity thought process

In the 1970s, acute narcosis toxicity was the model, short term LC50s were determined, and long term studies were mostly absent. Long-term protection was afforded by using a "safety factor". The document gives some history of the change in water quality standards for the state (TH to TaqH in 1996), but does not go into the change in the literature in recent years. Some of the key debate studies (DiToro, Barron) are covered in the annotated reviews, but the significance of the debate does not appear to be recognized. In setting water quality standard guidelines much of the literature from *Exxon Valdez* is of little direct use, as most of these studies were not controlled dosing studies. However, because of the growing literature on long- term effects from lingering exposure to chronic dose levels of oil, the *Exxon Valdez* studies do provide support for changing the existing toxicity model. While the concentration data for setting water quality standards is not available from these field studies, it is clear that there is a significant need to consider these sub-lethal effects when setting water quality standards.

The mechanisms of narcosis toxicity and chronic toxicity are fundamentally different. Acute narcosis toxicity, the basis for acute toxicity LC50s, is predominantly caused by one range of aromatic compounds (the mono and di-cyclic hydrocarbons), while long term chronic toxicity is likely caused by aromatic compounds (PAH), with some overlap. Narcosis is a quick acting mechanism, probably more of a membrane effect; homeostasis is rapidly affected, and if not immediately corrected, leads to mortality. Chronic damage however, is more intra-cellular and can be corrected, but costs energy. If there is enough damage, the fitness of the individual may be affected, and the organism lost through starvation or predation, rather than purely by a chemical death. Therefore, because the toxicity mechanisms are fundamentally different, the concept of applying a "safety factor" based on narcosis toxicity models may not necessarily safe guard against toxicants that act through different mechanisms. Some processes will be affected at different points in the life cycle of an organism, few of which are numerically significant at a point in time, but over time may accumulate and lead to losses within fishery stocks and subtle changes in the local ecology.

Water quality standards in the 1970s based on 4 day exposure tests need to be updated to take into account the chronic toxicity mechanisms, sensitive life stages, and sensitive life processes (e.g. reproduction including both physiological and behavioral aspects). The studies by Heintz et

al. where pink salmon embryos were exposed to oil, released at emergence to out migrate, avoid predation, capture prey, grow, sexually mature, and return back to their native stream accounts for a host of direct and indirect selection mechanisms that act on exposed individuals. Exposed embryos were less fit than their un-exposed cohorts. While studies of this size and magnitude are not practical or possible for most species, they do demonstrate the delayed chemical effects of sub-lethal doses.

Strategy employed

NMFS agrees that there are hundreds of papers to sort through, all with different species, life stages, exposure logistics, and analytical methods. However, because of the complex issues previously discussed, this is a critical stage in the review. Would "consensus agreement between species" or "consensus agreement between researchers" give the decision makers more confidence in their judgments?

Table 1, as the first example, compiles the ten lowest concentrations affecting Freshwater Fish. This table is dependent on three research papers, covering only two species, none of which are Alaska species. To correct this, NMFS recommends dividing Table 1 into two parts, with the five lowest concentrations reported for Non-Alaskan species and five for Alaskan species. If there were no Alaskan species, then the table would bring that to the readers' attention. Relying on one or two research papers is not appropriate. Additional research papers need to be represented. We suggest that the lowest concentration for any one paper be reported, rather than ten values from a single paper.

For example, while the strategy in Table 1 may be in the right direction, the specific information is in error. The table is designed for freshwater fish species. One paper represents eight observations, but the table is in error reporting the concentrations from this paper. The paper by Pollino and Holdway 2002, reports concentrations in mg/L according to the abstract, 3 orders of magnitude higher than the µg/L reported in the table. To further compound the troubles of Table 1, the other two citations are for a marine species (silversides), and the tests were conducted at 30 parts per thousand seawater. So, for Table 1, none of the data reported for freshwater species are correct, either in units or in test conditions. This raises issues on the validity of the remaining tables and points to the complexity of comparing data across species, researchers, units, etc.

Recommendation: Tables targeting the following, with verified concentration units, should be made. The lowest values from any one paper should be presented, not one paper for all listings.

Table 1. Acute toxicity in fish
Freshwater: lowest 5 Alaskan species and lowest 5 non-Alaskan species.
Seawater: Lowest 5 Alaskan Species and lowest 5 non-Alaskan species.
Table 2. Acute toxicity in invertebrates
Freshwater: Lowest 5 species
Seawater: Lowest 5 species
Table 3. Chronic or delayed toxicity
Fish

Embryos/Larvae of fish
Invertebrates
Embryos/larvae of invertebrates

Patterns would be expected to appear in this sort of comparison. The lowest values will likely show up in the more sensitive life stages. Long term tests, covering more sensitive life stages and processes, are more likely with embryos/larvae. Using different species/researchers, the patterns that emerge from the resultant tables will be more believable. In any case, tables should employee a single unit (μg/L) and convert all mg/L values to this. The tables should not mix total petroleum with total aromatics, total PAH (TPAH), or mix dispersed or UV enhanced hydrocarbons with straight hydrocarbons.

Specific comments

Relevancy of papers selected

There are hundreds of papers, particularly in the *Exxon Valdez* literature base that deal with the effects of oil, lingering oil, oil in sediments, and population effects. The following is a brief list of critical papers (post 1995) that have concentrations of PAH that should be considered in the evaluations of new water quality standards. These would be useful in understanding chronic toxicity mechanisms, persistence, and especially in providing a need for adequate water quality standards. The *Exxon Valdez* literature is the most complete suite of papers dealing with an acute spill, which over time evolves into a lingering chronic toxicity problem for a limited suite of species. That literature is not covered in the brief list of papers that follow.

Missing Post 1995 Acute Toxicity papers

Ara, K., K. Nokima, and J. Hiromi. 2002. Acute toxicity of Bunker A and C refined oils to the marine harpacticoid copepod Tigriopus japonicus. Bull. Environ. Contam. Toxicol. 69:104-100

Barron, M.G., T. Podrabsky, R.S. Ogle, J.E. Dugan, and R.W. Ricker. 1999. Sensitivity of the sand crab Emerita analoga to a weathered oil. Bull. Environ. Contam. Toxicol. 62:469-475

Gulec, I., B. Leonard, and D.A. Holdway. 1997. Oil and dispersed oil toxicity to amphipods and snails. Spill Sci. Technol. Bull. 4:1-6.

Long, S.M. and D.A. Holdway. 2002. Acute toxicity of crude and dispersed oil to Octopus pallidus (Hoyle, 1885) hatchlings. Water Res. 36:2769-2776/

Moles, A. 1998. Sensitivity of ten aquatic species to long-term crude oil exposure. Bull. Environ. Contam. Toxicol. 61:102-107.

Moles, A. 2001. Changing perspectives on oil toxicity evaluation. 2001 International Oil Spill Conference (Global Strategies for Prevention, Preparedness, Response, and Restoration). American Petroleum Institute Publication 14710, American Petroleum Institute, Washington DC.

Singer, M.M., S. George, I. Lee, S. Jacobson, L.L. weetman, G. Blondina, R.S. Tjeerdema, D,. Aurund, and M.L. Sowby. 1998. Effects of dispersant treatment on the acute aquatic toxicity of petroleum hydrocarbons. Archives Environ. Contam. Toxicol. 34:177-187.

Missing Post 1995 Chronic Toxicity papers

Carls, M. G., J. E. Hose, R. E. Thomas, and S.D. Rice. 2000. Exposure of Pacific herring to weathered crude oil: assessing effects on ova. Environ.Toxicol. Chem. 19:1649-1659.

Carls, M. G., G. D. Marty, T. R. Meyers, R. E. Thomas, and S. D. Rice. 1998. Expression of viral hemorrhagic septicemia virus in prespawning Pacific herring (*Clupea pallasi*) exposed to weathered crude oil. Can. J. Fish. Aquat. Sci. 55: 2300–2309.

Marty, G. D., J. W. Short, D.M. Dambach, N.H. Willits, R.A. Heintz, S.D. Rice, J.J. Stegeman, and D.E. Hinton. 1997. Ascites, premature emergence, increased gonadal cell apoptosis, and cytochrome P4501A induction in pink salmon larvae continuously exposed to oil-contaminated gravel during development. Canadian Journal of Zoology 75: 989-1007.

White, P.A., S. Robitaille, and J.B. Rasmusson. 1999. Heritable reproductive effects of benzo[a]pyrene on the fathead minnow (*Pimephales promelas*). Environ. Toxicol. Chem. 18:1843-1847.

Comments on Specific Tables

Page. 1, Table 1. The values reported by Pollino and Holdway are mg/L.

Comments on Specific Reviews

- Barron et al. 1999. Aquat. Tox. 46: 253

One important point missed by the reviewer is that heterocyclic aromatics (N,S,O analogs of polycyclic aromatics) are also important components in weathered oils and contribute to toxicity. The point of the paper is not that PAHs (ie, hydrocarbons) are not important in toxicity; rather it is not just PAHs, and except for dibenzothiophenes, heterocycles are not typically analyzed for.

- Brannon et al. 2001.

The review indicates that Brannon et al. does not link PAH exposure and mortality and that the paper offers little for TPAH water criteria standards. Perhaps this paper and others that fail to provide specific information should be considered in under the heading of "Additional References".

We suggest that ADEC consider replacing Brannon et al. 2001, with Thedinga et al. (in press) for a more balanced discussion of sample induced egg mortality, for the purpose of documenting potential sampling problems. In general, the hypothesis by Brannon et al. that mortality differences among streams was caused by differences in sampling relative to spawn timing is difficult to confirm or reject because daily run timing counts on individual streams were not practical (Rice et al. 2001). However, in 1991, the only year where adequate run timing data were collected, a statistically significant difference remained when sample timing was included as a covariate (Craig et al. 2002), leaving oil as the most likely reason for increased mortality. Furthermore, elevated embryo mortality in the field matched that in a controlled hatchery experiment with oiled and non-oiled eggs incubated under the same conditions (Bue et al. 1998).

- Bue et al. 1998

The Document concludes that this study provides little useful information to develop water-quality standards for hydrocarbon concentrations. The Document also questions the biological significance of mortality in oiled and non-oiled streams.

The 95% confidence bounds in Fig. 1 and 3 (Bue et al. 1998) are typically $\leq$20% and the mean mortality was consistently higher in oiled streams until 1994. We submit that the consistency of this pattern in leading towards recovery, the observation of elevated CYP1A in alevins in oiled streams (Wiedmer et al. 1996), the corroborative hatchery experiment (Bue et al. 1998), and experimental evidence of detrimental effects at low (part-per-billion) exposure concentrations together with elevated CYP1A (Marty et al. 1997; Heintz et al. 1999, 2000; Carls et al. submitted) collectively support the hypothesis of oil-induced damage.

- Carls et al. 1999

In paragraph two of the discussion on this paper, the reviewer suggests that Carls et al. (1999), were in error when concluding that exposure was to dissolved oil. Rather, the reviewer's opinion is that oil drops were formed by forcing water through oiled gravel. The reviewer's hypothesis is refuted by a recent overview of many oiled-rock column experiments and by comparison to a water accommodated experiment.

PAHs dissolve in water more readily than alkanes and the unresolved complex mixture. Thus, PAHs are enriched when oil is dissolved in water. Oil was dissolved in the majority of water passed through oiled rock columns. Percent total PAH [100 * TPAH/(TPAH + alkanes + UCM)] was centered at 43% and included 79% of the data. Percent TPAH in Alaska North Slope crude oil is 5.4% and declines with weathering. In contrast, only particulate oil was detected in water-accommodated fractions of crude oil prepared by mixing oil and water with a high-speed rotor (5.8% TPAH). Furthermore, Carls et al. (2002) strongly refute the claim by Pearson et al. (1985, 1995, 1999) that oil must adhere to herring eggs to cause negative effects (p. 165 - 166). The conclusion by Carls et al. that experimental herring were exposed to, and damaged by, dissolved oil is strongly supported by all available evidence.

In paragraph three of the discussion portion of this paper the reviewer questions how closely the composition of artificially weathered North Slope crude oil is to composition of *Exxon Valdez* oil (EVO) in the field. This weathering and further weathering as water passes through the oiled rock columns, exactly emulates the weathering process documented in Prince William Sound (Short and Heintz 1997). Carls et al. (2002) consider PAH composition in both field and lab and found good consistency between the two. Differences in weathering among treatment groups are quantified by reference to weathering factor *w*, estimated with the Short and Heintz (1997) model. None of the treatment concentrations merged; concentrations were distinctly different in all treatments at all times in both experiments.

The decision to treat concentrations below MDL, as discussed in paragraph four, is used so as to avoid erroneous "detection" of absent hydrocarbons. The reporting method, which is initial TPAH concentration, very likely overestimates adverse concentrations because aqueous concentrations decline exponentially in tests. The magnitude of this error is considerably larger than opposite errors potentially incurred by considering concentrations below MDL to be zero. The effect of MDL trimming is to slightly reduce the lowest concentrations (e.g., controls) with respect to other treatment concentrations. Dead larvae were typically necrotic and decomposing by the time they were recognized; there was no value in histopathological examination of these specimens.

The reviewer in paragraph five of the discussion correctly perceives that composition changed during exposure. This is expected in the lab and in the field. The most volatile (and least toxic) PAH disappear most quickly. The experiment clearly addresses the change in composition; witness comparisons between less weathered oil (LWO) and moderately weathered oil (MWO) exposures. Peak tissue concentration reported in the LWO experiment was just that - the maximum reached under experimental conditions, not the theoretical peak. Heintz et al. (1999) provide a model to estimate peak concentration as a function of changing PAH concentration in water and in tissue.

In paragraph six, the reviewer expresses concern that temperature differences confounded the estimate of reduced hatch time. Temperature was not a confounding factor within experiments and the dose-effect was clear and consistent in both. Initial TPAH concentrations were conservatively used to judge reactions. A good alternative is geometric mean aqueous TPAH concentration; this model accounts for time and the exponentially decaying concentrations. Geometric mean concentrations ranged from 0.06 (control) to 40 μg/L in the LWO and 0.03 to 2.1 μg/L in the MWO. An even better integrator of exposure concentration and time is egg tissue; this relationship is demonstrated in Barron et al. (2004). Dose responses are clear regardless of choice of independent variable. Barron et al. (2004) found that narcotic toxicity (acute toxicity) was possible only in the highest dose the herring assays.

The suggestion in paragraph seven of the discussion, that embryos in the MWO experiment were exposed to a higher time concentration is not possible; concentrations declined in all cases and exposure times were identical. Furthermore, tissue uptake clearly indicated a large time-integrated difference in exposure concentration. The conclusion that MWO was more toxic holds regardless of choice of independent variable, including initial aqueous concentration,

geometric mean aqueous concentration, or PAH concentration in tissue. Recent evidence suggests naphthalenes were of little importance in these tests and other work suggests that the more environmentally persistent PAH might be the most damaging. Median lethal concentrations based on persistent PAH were very similar between less- and more-weathered tests, suggesting these are the primary toxins.

In table 11 the entry for the larval mortality lowest observed effects concentration (LOEC) is incorrect and should read 22 to 108 ng/g wet weight (total PAH in tissue); this corresponds to 0.4 to 0.7 μg/L (aqueous total PAH concentration). Also, values in table 11 are marked "≥" but should be marked "≤" because the values reported were initial aqueous concentrations.

In summary, the reviewer suggests this paper has limited utility in helping establish water quality criteria. We respectfully disagree. This paper, along with a number of similar salmon papers, provides the best evidence ever published that stricter PAH standards are needed. The comprehensive data of this study allow in depth analyses to address the reviewer's suggestion to account for both time and concentration.

We submit that the reviewer misses several key points (1) that use of initial aqueous concentrations (as published) yields a very conservative estimate of PAH toxicity; (2) that geometric mean concentrations are estimable and account for time; and (3) that PAH concentrations in tissue provide a time-weighted summary of exposure. The conclusion that MWO was more toxic than LWO holds regardless of the choice of independent variable; including initial aqueous concentration, geometric mean aqueous concentration, or PAH concentration in tissue and is consistent with a large body of literature that finds increasing toxicity with molecular size and alkyl substitution. The suggestion that exposure was due to oil droplets is unfounded; the data demonstrate quite the opposite. Short and Heintz (1997) demonstrate with great rigor that the methods used in this experiment emulate Exxon Valdez hydrocarbons in Prince William Sound.

The reviewer is correct that rapidly falling concentrations complicate application of the results to water quality criteria. However, the most rapidly changing constituents (naphthalenes) likely contributed least to toxicity. In contrast, environmentally persistent PAH were essentially constant in the upper doses. A comprehensive amalgam of studies from scientist at NMFS' ABL provides ample evidence for stricter standards, even when the most conservative concentration metrics are considered. We are willing to work with regulators to fully supply data and develop the analyses required tailored to regulatory needs.

- Carls et al. 2002

In paragraph two of the discussion on this paper, the reviewer cites Short & Heintz (1997). This is an incorrect citation for this information. The correct citation for this information should be Carls et al. (1999). Please note that the data in table 12 originate from the same 1999 study as the data in table 11.

Heintz et al. 1999

The reviewer indicates in this and other NOAA papers, that the reported Method Detection Limits (MDL) are too high and the statistical effects of censoring values below MDL is questioned. We are not clear on how MDLs less than eight parts per trillion (Short et al. 1996) can be construed as too high. The MDLs were set at the levels at which we were 99% certain that the signal differed from background noise in accordance with the Environmental Protection Agency's "Guidelines for Establishing Test Procedures for the Analysis of Pollutants; Procedures for Detection and Quantitation,"(Appendix B, 40 CFR Part 136). Values below MDL were censored in order to minimize the chances of ascribing effects to non-existent PAHs. This approach has the potential of reducing the lowest effects concentration (LOEC) when PAHs are considered in aggregate. A frequently described alternative is to replace values below MDL with the MDL value, thereby minimizing the potential for underestimating the LOEC. This approach has a trivial effect on the data presented in Heintz et al. (1999), increasing the initial aqueous exposure from the very weathered oil dose from 1.0 to 1.0 parts-per-billion. As to the statistical effect of censoring noise, none of these reported summed PAH values were used in statistical analyses in any way other than as ordinal factors in the Analysis of Variance (ANOVA). Consequently, censoring the reported concentrations below MDL has no effect on the outcomes of the tests.

Marty, G.D., Okihiro, M.S., Brown, E.D., Hanes, D., Hinton, D.E., 1999. Histopathology of adult Pacific herring in Prince William Sound, Alaska, after the Exxon Valdez oil spill. Canadian Journal of Fisheries and Aquatic Sciences 56, 419-426.

On pages 39 and 97 of the Document, the year for this article is incorrectly listed as 1997.

Statement from Document

> "Results of concentrations of naphthalenes in fish viscera showed no difference between reference and oiled samples in 1989; however, in 1990, naphthalene was higher in many samples for the "reference" animals."

Page 422 (column 2, near top) of the Marty paper says, "Tissue concentrations of TPAH were significantly greater in fish collected from oiled areas in 1989 than in fish collected from reference areas in 1989. Naphthalenes represented 95-97% of TPAH above method detection limits in these tissues..."

Page 424 (Table 7, footnote) of the original paper states: "None of the site differences were statistically significant."

Suggested revision to statement

> "Concentrations of TPAH in fish viscera were significantly greater in fish from oiled sites than in fish from reference sites in 1989; however, in 1990, concentrations of aromatics and naphthalenes in fish viscera showed no difference between reference and oiled samples."

Statement from Document:

> "The discrepancy between 1989 and 1990 results is likely due to limited sample size."

This statement is speculative and is an incorrect interpretation. Without additional studies to support it, we recommend it be deleted. This paper offers two hypotheses for differences in results from 1989 and 1990 that are supported by the data.

Statement from Document:

> "The report's conclusions that differences in pathology resulted from differences in concentrations of hydrocarbons is not supported by the results and should be viewed as speculative."

The last sentence of the Abstract for this paper states that: "It is concluded that Pacific herring were exposed to Exxon Valdez oil in 1989, and that development of hepatic necrosis in exposed fish probably was a result of VHSV expression."

Suggested revision to the Document (based on last paragraph of discussion on page 425 of the paper)

> "The report's conclusion that Pacific herring are unusually sensitive to effects of crude oil if exposure occurs in the weeks before spawning are based on experimental evidence (see Carls et al. 1998)."

Statement from Document:

> "Further, no dose-response curve was assessed so it is not possible to relate observed effects to exposure concentrations."

We agree. However, lack of dose-response curves is standard methodology in field studies of actual oil spills. Field studies are best used to validate laboratory experiments. Therefore, results in this paper need to be compared with Carls et al. 1998, as suggested by Marty et al.

Statement from Document:

> "The opening line of the Abstract [Pacific herring (Clupea pallasi) sampled from oiled sites in Prince William Sound, Alaska, U.S.A., 3 weeks after the 1989 Exxon Valdez oil spill had multifocal necrosis and significantly increased tissue concentrations of polynuclear aromatic hydrocarbons (PAH)] is misleading and not supported by the results presented in this report."

This statement is an incorrect interpretation. We recommend that the reviewer reread the statistical analyses along with Carl et al. 1998. We recommend this statement be deleted.

• Little et al. 2000

An important point not brought out regarding this paper is that it also demonstrated that total PAHs can be toxic below 10 μg/L in the presence of UV.

•Wilson et al. 2000

We agree with the reviewer that this paper has limited value in supporting water quality criteria.

Additional Review Comments

The following provides additional information to that of the reviewer for specific papers included in the Document.

•Neff et al. 2000

Table 7 of Neff et al. is entitled "Acute Toxicity" but is actually percent water accommodated fractions (WAF). The percentages in this table must be multiplied by the concentrations listed in other tables to determine 96 h concentrations. When this is completed, a consistent pattern emerges; toxicity increases with weathering for all species and in all oils. We suggest replacing the %WAF column in table 19 with total hydrocarbon concentration. Also, there is an error in the third record in table 19. The %WAF is 48%, not 4%. In addition, we suggest the following information be inserted into table 19.

LC50/EC50s expressed as mg/L total hydrocarbons

		clownfish	**silverside**	**mysid**	**shrimp**	**urchin**	**Echino**
Wonnich	crude	12.69	11.61	17.41	10.88	15.96	3.99
	150*	1.50	1.72	0.94	0.83	0.81	0.52
	200		3.97	1.76	1.93	1.71	3.36
Campbell	condens	14.94	14.18	13.79	11.49	26.05	12.64
	150	11.36	12.49	5.33	5.33	11.22	
	200			1.02	0.96		
Agincourt	crude*		0.01		0.01		
	150			0.10			
	200			0.13			
	250			0.01			
Diesel	fresh		0.65	0.36	0.37		0.32
	200		0.67	0.27	0.26		
	250		0.40	0.19	0.20		0.32

*mono-aromatic hydrocarbons were not measured for this weathering treatment.

Relating the experiment by Neff et al. (2000) to environmental conditions is difficult. Weathering was accomplished by heating the oils to high temperatures (250°C), driving off the lightest compounds. Unlike the weathering of spilled oil, where the absolute PAH concentration declines, absolute PAH concentrations increased with weathering in this experiment. This artifact may partially explain why toxicity appears to decrease with weathering when expressed as total PAH.

However, it is possible that these bioassays were insufficient to fully measure PAH toxicity. The assays were designed to investigate short-term acute responses only, missing slower-acting non-narcotic damage caused by PAH. Nonetheless, the hazard analysis provided in table 10 of Neff et al. (2000) indicated that PAH explained an increasing proportion of toxicity as weathering increased and was the dominant factor in oils heated to ≥200°C.

Comparison of this paper to others demonstrates that how assays are conducted, as well as composition of oil, is of great importance in determining water quality standards. We agree with Zhao & Newman (2004), that short-term, acute bioassays have limited relevance where toxic response is slow and measurement of damage requires longer periods of time. For example, Heintz et al. (2000) report diminished adult pink salmon returns as a result of embryonic exposure to oil, a lifetime impact. Where oil was weathered in an environmentally realistic way, total PAH concentrations declined, yet relative concentrations of environmentally persistent PAH increased and PAH toxicity increased (e.g., Carls et al. 1999; Heintz et al. 1999). Acute toxicity caused by monoaromatic compounds was not a significant issue after the *Exxon Valdez* oil spill. One-ring compounds evaporated within a few days (e.g., Wolf et al. 1994; Neff & Stubblefield 1995), thus reports of remaining oil and oil toxicity in Prince William Sound focus on PAH toxicity. Long-term chronic assays are typically required to fully appreciate the non-narcotic toxicity of PAH, hence the work by Neff et al. (2000) should be considered a starting point. An in depth discussion of assay methods is needed to place toxicity measures on an even footing and avoid inappropriate use of toxicity estimates.

• Neff 2002

In this paper Neff adopts the opinion of Bongiovanni et al (1989) and Ehrhardt et al. (1992) that photooxidation and biodegradation are the most important weathering processes for PAHs in the marine environment. This statement misses evaporation (which Neff covers separately) and dissolution. Extensive weathering is possible via dissolution, a process clearly dominant in laboratory studies (e.g., Marty et al. 1997; Carls et al. 1999; Heintz et al. 1999, 2000) and apparent in Prince William Sound sediment.

On page 302 of Neff (2002), data concerning Prince William Sound are in error. Short & Harris (1996) report TPAH concentrations as high as 45.2 μg/g, not 6.3 μg/g as stated by Neff. Similarly, Neff states TPAH in mussels on shoreline exceeded 40 μg/g – we agree, but the value reached 266 μg/g. The cited water quality criteria (Table 107) are likely BTEX acute lethal values, and of questionable use in setting PAH standards. The ERL and ERM (NOAA's Effects Range Low and Effects Range Medium). estimates in table 108 should be reviewed with respect

to aqueous transfer to organisms. For example, an original oiling of 2860 ppm, weathered to 4.6 ppm over a one-year period produced an initial aqueous concentration of 1 ppb; pink salmon embryos were damaged by exposure to this contaminated water (Heintz et al. 1999).

Neff discounts modes of toxicity other than narcosis for 2 to 4 ring PAH. This is inconsistent with his own observations that toxicity increases with molecular size and alkylation. Although Neff recognizes that "hydrocarbons in solution are the most bioavailable and toxic to marine organisms," but does not present the long-term assays necessary to demonstrate the non-narcotic action of PAH and consistently finds fault with such tests that provide this evidence.

Page 309 of Neff (2002) discusses tests that demonstrate microbial action is required for water-accommodated fractions of oil to be toxic. The cited tests have a fatal confounding variable; aqueous PAH concentrations increased which microbial action. Bacterial surfactants promoted entry of oil constituents into water, and these constituents were toxic.

Neff's statements that microbial growth explains toxicity in oiled-rock column tests (e.g., Carls et al. 1999; Heintz et al. 1999) are inconsistent with the biological and chemical data. Carls (2002) provides a discussion of microbial metabolism and toxicity and rebuts the Neff hypothesis, concluding that metabolic byproducts accounted for <3% of observed toxicity in laboratory tests and, by inference, contributed little to observed toxicity in PWS. Comparison of NRDA results, which demonstrate the presence of PAH from oil in eggs, with the pure compound tests by Incardona et al. (2004) clearly indicates that these unaltered PAH are toxic and produce the same suite of embryonic responses.

• Pollino and Holdway 2002

This paper reports the results of acute, static-renewal exposure of a freshwater fish to crude oil, dispersants, and naphthalene. Embryonic response to oil is consistent with other larval research, including edema, spinal deformities, and abnormal jaws. Oil preparation methods are vague and no attempt was made to characterize hydrocarbon composition in the exposures. Abnormalities were observed at ≥0.5 mg/L total petroleum hydrocarbons. The authors suspect that PAH may account for most of the toxicity and suggest that the oil may contain about 1.5% PAH, thus rainbowfish may be negatively affected by <10 µg/L TPAH. This response is consistent with recent work with Pacific herring, (Carls et al. 1999) and pink salmon, (Marty et al. 1997; Heintz et al. 1999 and 2000). The results on rainbowfish are also consistent with observations of another freshwater fish, zebrafish, where abnormalities were observed at roughly 50 µg/L (unpublished data).

• Roe-Utvik & Johnsen 1999

This is the only paper included in the Document that addresses Passive Sampling Devices (PSD). We recommend a comprehensive review of PSDs. There are many published studies, including contributions from our lab (Carls et al. 2004). We agree with the reviewer that membrane-based PSDs can be used to collect hydrocarbon data (Carls et al. 2004). However, passive and living samplers do not always provide the same answer (e.g., Carls et al. 2004). We recommend

detailed parallel study before any PSD is used alone as a substitute for a living organism. The best use of PSDs may be to augment other data, sample hydrocarbons where concentrations are low and or sporadic, or simply identify problem areas for further study.

Roe-Utvik & Johnsen demonstrate similar PAH accumulation in semi-permeable membrane devices (SPMDs) and mussels grouped together along a known hydrocarbon gradient. Similarity in PAH composition between the passive and living matrices suggests exposure of each to an identical source, yet theoretically SPMDs sample only dissolved PAH and mussels are capable of sampling both dissolved and particulate oil. ABL field experience suggests that the route of exposure in mussels (from Prince William Sound) to oil is principally but not entirely particulate.

The conclusion by Roe-Utvik & Johnsen that water-soluble fractions are the most important route of exposure for marine organisms at lower trophic levels is based on the apparently unpublished work of colleagues and is not defensible from this work alone. While we agree that dissolved hydrocarbons are the cause of toxicity in living organisms, such as developing fish embryos, particulate oil in water has the same effect because soluble constituents dissolve into the water. We recommend additional discussion of this topic in the Document.

Comments on Pre-1995 Research

This section was limited by intent to Alaskan species or PAH research that might be particularly relevant to understanding PAH toxicity. The author of this section synthesizes the state of pre-1995 knowledge. Heavy weight is placed on Nagpal's tables, which mix pure compounds, PAH mixtures, static and flow through tests, and measured and unmeasured concentrations. Without the synthesis, the table would be overwhelming and of far less value. A heavy emphasis is placed on the acute toxicity results of Rice et al. 1976, perhaps the most out-of-date work available on crude oil toxicity to Alaskan species. Rice et al. (1984), summarizes both the lethal toxicity and sublethal effects of exposure to crude oil water-soluble fractions for a large number of studies done with Alaskan species. However, the following references still remain the most comprehensive suite of acute toxicity work done on Alaskan species with BTEX (benzene, toluene, ethylbenzene, and xylene heavy crude oils).

Rice et al. (1979) "Sensitivity of 39 Alaskan marine species to Cook Inlet crude oil and No. 2 fuel oil." *In* API, EPA, and USCG, 1979 Oil Spill Conference (Prevention, Behavior, Control, Cleanup), pp. 549-554. Proceedings of a symposium. American Petroleum Institute, Washington, D.C.

Moles (1998) "Sensitivity of ten aquatic species to long-term crude oil exposure." Bull. Environ. Contam. Toxicol. 61:102–107.