UNITED STATES ENVIRONMENTAL PROTECTION AGENCY
WASHINGTON, D.C. 20460

April 17, 2009

OFFICE OF
WATER

Stephen F. Sundlof, D.V.M , Director
Center for Food Safety & Applied Nutrition
U.S. Food and Drug Administration
5100 Paint Branch Parkway
College Park, Maryland 20740-3835

Re: Comments on the Draft FDA Report Assessing Risks and Benefits from Fish

Dear Dr. Sundlof:

Enclosed is a copy of the comments that the Environmental Protection Agency (EPA) is submitting to the docket in response to the notice in the Federal Register on January 21, 2009 (74 FR3615) announcing the availability of two draft documents: 1) "Report of Quantitative Risk and Benefit Assessment of Commercial Fish Consumption, Focusing on Fetal Neurodevelopmental Effects (Measured by Verbal Development in Children) and on Coronary Heart Disease and Stroke in the General Population" [draft risk and benefit assessment report]; and 2) ``Summary of Published Research on the Beneficial Effects of Fish Consumption and Omega-3 Fatty Acids for Certain Neurodevelopmental and Cardiovascular Endpoints" [draft summary of published research].

EPA has serious fundamental concerns on the underlying science and methodology used by FDA. Our concerns were previously communicated to FDA in the fall/winter 2008 during the interagency review of a draft of the reports, prior to the release of these reports on January 21, 2009, for public comment. The attached comments are from EPA scientists and others who provided a thorough and thoughtful review of the FDA documents. The comments are submitted in the interest of furthering the goals of achieving the highest level of scientific integrity in the analysis of scientific information which will ultimately be used in protecting public health.

I am certain that you share these goals and will continue to work with EPA in furthering the public interest. If you have any questions or concerns, you may contact me at (202) 566-1566, or keehner.denise@epa.gov.

Sincerely,

Denise Keehner

Denise Keehner, Director
Standards and Health Protection Division
Office of Science and Technology

Enclosure

cc: Dr. Randall Lutter
Deputy Commissioner, Office of Policy, Planning & Preparedness (FDA)

Mr. Michael Shapiro
Acting Assistant Administrator, Office of Water (EPA)

Mr. Ephraim King
Director, Office of Science and Technology (EPA)

Dr. Rita Schoeny
Senior Science Advisor, Office of Water (EPA)

Mr. Jeff Bigler
Fish Advisory Program Lead, Fish, Shellfish, Beach and Outreach Branch (EPA)

Mr. John Wathen
Assistant Chief, Fish, Shellfish, Beach and Outreach Branch (EPA)

Ms. Denise Hawkins
Chief, Fish, Shellfish, Beach and Outreach Branch (EPA)

04 17 09

U.S. EPA PUBLIC COMMENTS:

"Report of Quantitative Risk and Benefit Assessment of Commercial Fish Consumption, Focusing on Fetal Neurodevelopmental Effects (Measured by Verbal Development in Children) and on Coronary Heart Disease and Stroke in the General Population''; and 2) ``Summary of Published Research on the Beneficial Effects of Fish Consumption and Omega-3 Fatty Acids for Certain Neurodevelopmental and Cardiovascular Endpoints''

BACKGROUND

U.S. EPA is providing the following comments in response to the U.S. Food and Drug Administration (FDA) Federal Register Notice of January 21, 2009 (Vol 74, Number 12, pp 3615-3617. This notice announced availability of two draft documents: 1) "Report of Quantitative Risk and Benefit Assessment of Commercial Fish Consumption, Focusing on Fetal Neurodevelopmental Effects (Measured by Verbal Development in Children) and on Coronary Heart Disease and Stroke in the General Population" [Risk / Benefit Report]; and 2) ``Summary of Published Research on the Beneficial Effects of Fish Consumption and Omega-3 Fatty Acids for Certain Neurodevelopmental and Cardiovascular Endpoints" [Omega-3 Report].

EPA provided written comments on an earlier draft FDA analysis through a formal Interagency Review process facilitated by Office of Management and Budget (OMB). The last set of comments were provided by EPA on December 18, 2009. EPA is not in receipt of a formal response to Interagency comments.

OVERVIEW OF EPA COMMENTS

1. EPA agrees that a comprehensive analysis of risks and benefits of fish consumption is important to supporting informed public health decisions. EPA also believes that in an era of competing needs in every Agency, that the priority of such an analysis should be agreed upon up front by all involved Agencies.

The conceptual approach underlying the draft analysis reflects innovative methods development. EPA feels that the draft analysis provides a valuable starting point for development of a future comprehensive integrated exposure and risk assessment that could inform outreach and policy initiatives on fish consumption and public health.

EPA feels that such an analysis would best consider multiple contaminants as well as various components of fish that may confer benefits to consumers. This should be a joint effort among involved Agencies including but not restricted to FDA, EPA, Department of Health and Human Services (DHHS), and Department of Commerce (National Oceanic and Atmospheric Administration, NOAA). An appropriate vehicle for a risk / benefit analysis may be a Memorandum of Understanding that specifies roles for the Agencies as well as resources to be used in the effort. Given the expected large allocation of resources for a comprehensive analysis of a complex situation, the priority for this effort must be agreed upon by all the involved Agencies.

2. **<u>An analysis with significant implications for public health policy, such as this one, should include a formal design phase that is subject to peer review.</u>**

The high-profile nature of this topic, as well as the innovative nature of the integrated approach to simultaneous consideration of fish risks and benefits, argues for a systematic approach to designing the analysis. This systematic design phase includes the use of peer-review. Such an approach would first identify the list of policy-related questions being addressed by the modeling effort.[1]. Once key policy-related questions are identified, then the specific set of risk and benefit metrics that would ideally be generated to address those questions can be specified. The third step in model design would be to identify conceptual modeling approaches for generating these risk and benefit metrics. With a topic as complex as that addressed in the draft FDA analysis, it is likely that several conceptual approaches for modeling risks and benefits would be identified (each of these conceptual approaches, in turn, likely having different specifications of models and datasets). This design would require rigorous peer review and revision of the study plan before the joint Agency efforts would commence.

EPA feels that for en effort of this magnitude and public health significance, the absence of a systematic, collaborative, peer review design phase makes any risk-benefit analysis vulnerable to challenge from stake-holders; moreover, poorly designed risk-benefit analyses can lead to stakeholders viewing the entire risk-benefit paradigm as flawed. EPA feels that the current FDA Risk / Benefit Report suffers greatly from this lack of collaborative planning and the transparency this brings to the entire process.

3. **<u>Peer review of the draft FDA analysis was inadequate and did not follow the OMB bulletin ("Final Information Quality Bulletin for Peer Review").</u>**

The Risk / Benefit Report meets criteria for a Highly Influential Scientific Assessment as described in the OMB Bulletin (e.g. "is novel, controversial or precedent setting or has significant interagency interest"). Generally a letter review by a limited number of reviewers is not considered adequate for a report of this importance; a panel review that is open to the public is recommended. It should also be noted that the expertise of the particular reviewers used did not extend to the full range of material in the Risk/ Benefit report.

EPA did participate in an Interagency review of an earlier draft document. The Agency acknowledges and appreciates the changes made in the current draft to reflect some of these comments. EPA feels that the Risk / Benefit Report remains a good proof of

[1] Although the draft FDA report clearly states that it is not generating risk-benefit results reflecting any form of risk management decision, it is difficult to make the case that the results of this report would not be interpreted as having implications (either directly or indirectly) for public health policy decision making (e.g., outreach messages to the public regarding fish consumption). Therefore, in designing an analysis such as this, identifying the specific policy-relevant questions likely being addressed (or for which the results of the analysis might be applied), would be a first step to undertaking model development.

concept and could serve as the basis for an assessment of data needs and analytic development

4. **While some changes have been made in response to EPA comments on the earlier draft, the analyses themselves are essentially unchanged, and in the opinion of EPA, scientifically flawed. The current draft of the Risk / Benefit Report is not an appropriate basis for public policy decisions. Moreover, as stated above, EPA believes that an appropriate analysis requires a thoroughly critiqued design phase.**

The deficits in the draft Risk / Benefit Report do not allow for simple revision, but rather that the analytic plan should be thoroughly re-thought. Some of the most significant problems include these.

1. The choice of critical study and endpoint for the neurodevelopmental assessment is critically flawed and these choices fundamentally affect the conclusions drawn from the assessment with regard to the balance between mercury risk and fish nutritional benefits. Different study choices and different endpoints would likely fundamentally change the overall conclusions reached.

2. There are major deficiencies in the literature cited, and many papers described are not accurately characterized.

3. Model descriptions are abstruse and incomplete to the extent that it is difficult to fully evaluate the scientific validity of either the methods or the results. Many assumptions described have insufficient rationale for their choice.

4. There are serious concerns with statistical methods, in particular the derivation and use of Z-scores. (A "z score" is a common statistical way of standardizing data on the same scale so that comparisons can be made.)

5. The analyses do not adequately characterize the populations at risk.

6. There are substantial difficulties with the analysis of cardiovascular risk, including the failure to analyze non-fatal coronary heart disease.

In the next section EPA provides specific comments on both the draft Risk / Benefit Report and (in an addendum) comments on the draft Omega-3 report.

EPA COMMENTS ON THE JANUARY 2009 DRAFT FDA RISK / BENEFIT REPORT

1. General Comments

The population at risk is not appropriately defined or evaluated.

Vulnerable populations were not adequately discussed

There was no discussion of vulnerable populations in the draft analysis, in particular populations exhibiting high fish consumption rates or populations consuming large quantities of particular fish species.

Since the FDA document deals with fish consumption as a fixed average composite diet with respect to fish species/type and the frequency of consumption of each species/type, the draft analysis did not address populations and/or individuals who are vulnerable because their patterns of fish consumption significantly depart from the national composite average. This is a major problem with the draft analysis. Mahaffey et al. (2009) have clearly shown that such variability is quite significant on a regional basis. Case reports of consumers with high levels of mercury in biological samples makes it clear that such variability certainly exists among individuals even within regional patterns. In addition, the FDA modeling does not address toxicokinetic variability in the population (for either fetal or maternal dose). The National Academy of Sciences (NAS), (NRC 2001) recommended that pharmacokinetic variability be factored into any risk assessment for methylmercury (MeHg).

"Unit fish" or "fish as a package" severely impairs the predictive nature of model results.

A major and fundamental conceptual problem with the draft analysis is that it views fish consumption as consisting of a fixed average diet across the population. That is, it assumes that under any given scenario, fish is a homogeneous mixture of that presents the same concentration of omega-3s (and other beneficial nutrients) and MeHg to all consumers and that consumer differ in their fish consumption only in the amount of fish they consume. Clearly, however, different consumers consume different types and species of fish with different frequencies. These different types and species each have characteristic levels of omega-3s and MeHg. The approach used by FDA facilitates a very complex modeling design, but is unable to distinguish different dietary patterns among consumers that result in different combinations of omega-3s and MeHg, and thus, different ratios of risk and benefit. Fish consumption advice based on the FDA approach will result in both unintended risk and reduced benefit. This is particularly the case for consideration of cardiovascular effects. There, the does not even consider the MeHg-specific risks, but assumes that the fish "package" includes a fixed MeHg component.

A closely related problem is that since fish is considered to be a homogeneous mixture, the what-if scenarios that investigate the impact of increasing of decreasing fish consumption, only address changes in the *amount* of fish consumed, but implicitly assume changes in fish consumption will not affect the types of fish that are consumed. This assumption is not validated and appears to be highly unrealistic. Changes in consumption tend to follow trends in the popularity of specific market items rather than necessarily changing consumption across the board. Good examples of this are the large increases in the popularity of salmon and sushi tuna in recent years.

The rationale for this highly reductive approach is the underlying structural assumption (also present in the first draft but apparently first acknowledged in the current draft), that "The assessment is intended to be nationally representative of the U.S. population. It does not address risk to segments of the population whose exposure to MeHg or patterns of fish

consumption may be substantially different from the population as a whole (pg. 7)." But that segment of the population, however one defines it precisely, is likely to be the at-risk population. It is known that 5-10% of the population (depending on the specific region) of women of childbearing age exceeds a health benchmark, the EPA's MeHg reference dose (RfD). The at-risk population may, in fact, include some fraction of the population whose MeHg exposure is below the RfD. Clearly, however, the fraction of the population that exceeds the RfD is likely to be the segment of the segment of the population "whose exposure to MeHg or patterns of fish consumption may be substantially different from the population as a whole." One would *a priori* assume that a risk-benefit analysis would, at least, address itself to that segment of the population that is likely to be most at risk. Even if that segment of the population was a miniscule percent of the whole, one would assume that this would be established before proceeding to focus the remainder of the analysis on a different population group. In this case, however, with reference to the RfD, it appears likely that the at-risk population that is excluded from the FDA draft analysis is quite substantial. By addressing the fraction of the population that is most likely not to be at risk to begin with, the analysis becomes *de facto* a benefits analysis rather than a risk-benefits analysis. This characterization is reinforced by the assumption of a threshold for MeHg adverse effects but no ceiling for fish's beneficial effects, as was used for the neurodevelopmental model.

Consequently, consumer advice about fish *per se* that might not place the average fish consumer at risk and might provide overall benefit, might create significant risk for the at-risk population (or the population that would move into the at-risk category as a result of this generic fish consumption advice). Viewed in this context, a fish consumption advisory strategy based on the design of the FDA draft analysis would be highly inconsistent with what is generally considered to be proper public health practice.

It is not clear what population or percentage of the U.S. population is covered in the risk / benefit analysis. It appears that the advantages of fish consumption are calculated for some average consumer.

A key decision to be made in a systematic design of a risk / benefit analysis is whether the model is intended to cover risk-benefit tradeoffs for specific subpopulations of the commercial fish consuming population. For example, there are segments of the U.S. population whose consumption of commercial fish is not typical, but who may still represent relatively large population groups and who importantly, may occupy positions higher-up on a distribution of U.S. population impacts/benefits from commercial fish consumption. The draft Risk / Benefit report is not clear in specifying whether the model is intended to cover these key sub-populations. However, it does provide estimates for the 10^{th} to 99^{th} percentiles of the commercial fish consuming population, which would likely include some of these subpopulations. Clarity in terms of the subpopulations to be covered by a modeling effort is a key first step in approaching the design of a modeling approach.

Key (commercial fish consuming) subpopulations were not included in the draft analysis.

It would be most informative to include risk-benefit tradeoff estimates for a set of clearly-defined subpopulations of commercial fish consumers. These estimates could be either

deterministic or probabilistic in nature, but would involve modeling the risk-benefit tradeoff for some number of viable (plausible) subpopulations, which likely occupy points higher up on the overall U.S. population risk-benefit distribution generated using the FDA model. These could include, for example, wealthy consumers who eat several higher trophic level fish meals a week, such as swordfish, or poorer consumers who eat several meals a week of the less expensive commercial fish. The point of including these discrete scenarios would be two-fold. First, they would provide some focused assessments on subpopulations of concern; giving us specific risk-benefit tradeoff estimates for these groups; presumably this would be of great interest to the public. And second, they could be used to support performance evaluation of the FDA's larger national-scale consumer model. Specifically, the results of these individual sub-population modeling efforts could be compared to the larger national distribution to see whether the relationship between these two types of estimates seems "reasonable" or conversely, if they are so discrepant that concerns are raised about high-end consumer modeling in general, using the FDA model. It is recognized there are concerns in conducting this type of special subgroup modeling, in that one would not know how representative they actually are of the broader national population (i.e., how many of each sub-population really there are in the U.S.). Ideally, these sub-populations would have been defined such that there is some degree of confidence that they represent a real (if small) segment of the broader consuming population.

The document lacks transparency in many places.

Some examples are these.

1. It is difficult to assess the meaning of a comparison (e.g., between exposure levels for different populations) in the absence of any reference values. The document would be much clearer if the actual values were provided. For example:
    a. On p. 11, it is stated that US women of childbearing age are exposed to 1/15 the level of methyl mercury of Seychelles women and 1/10 that of Faroe Islands women, with no exposure levels provided for either group;
    b. On p. 21, it states that women in the Seychelles eat 12 fish meals per week to achieve body levels 10-fold higher than US levels, again without referencing any actual body burdens.
2. In numerous places, the document presents conclusions without clearly stating the assumptions behind them, or in ways that obscure the assumptions. The document now includes a table summarizing studies and findings (p. 23 Table IIIA). Some of the findings are not clearly presented, which could lead a reader to misunderstand the results; see page-specific comments in subsequent section.
3. The revised document includes a table listing assumptions and limitations related to the choices made for the dose-response modeling, and describing the impact of these assumptions on the model estimates. Inclusion of this table provides a definite improvement in the document, and increases the clarity of the choices made by the authors in modeling these data. However, we found a number of the listed assumptions to be unfounded, and in some cases the potential impact of these assumptions on the model results was understated or confusing.

See comments by page number for additional specifics.

Terms used throughout the document should be clearly defined.

Terms such as "background levels of Hg" and "high methylmercury-to-fish ratios**"** should be defined carefully and various statements of concentrations (i.e., <0.1 ppm) be replaced in the analysis with the actual concentrations.

2. There are serious concerns with statistical methodologies and several aspects of the modeling.

There are serious concerns about use of z-scores in this draft analysis, both with respect to the ways in which they were derived and in the way they were used to compare across endpoints.

The derivations of the z-scores used are questionable.

As described in section IV of the document (e.g., p. 116), and discussed at the OMB-convened EPA/FDA meeting on 12/2/08, the standard deviation (s.d.) from the Seychelles data on age of talking was used to derive the z-scores for this measure. It was repeatedly stated that the slope of this model was driven mostly by the (high dose) Iraqi data; thus, the appropriateness of using a z-score derived from the Seychelles data is questionable. Since (1) the size of the z-score depends entirely on the s.d., and (2) the argument supporting the validity of using these 'walking and talking' data to represent neurobehavioral effects of methylmercury (MeHg) is based (at least in part) on a comparison of the "age at talking" z-score to the IQ z-scores derived from other studies, inappropriate selection of the data (i.e., the s.d.) used to derive the z-score could undermine the validity of this draft analysis . This is even more critical since the benefit/risk draft analysis is apparently also based on combining z-scores from the "age at talking" analysis with z-scores based on an analysis of data from the Daniels et al (2004) study. [Note also that the s.d. used to derive the z-scores for the "age at walking" data was not provided.] The EPA/FDA meeting on 12/2/08 resulted in a consensus that a rationale for selection of the s.d. used in calculating the z-scores should be provided, along with a discussion of the sensitivity of the conclusions to this choice. It was suggested that a sensitivity analysis be conducted using alternative s.d. values (e.g., from the Iraqi data or from the U.S. population).

The draft analysis uses z-scores in an inappropriate way.

The draft analysis makes "apples and oranges" comparisons without considering biological context. It is unclear whether a z-score calculated using raw "age at talking" data is appropriately compared to a z-score calculated on the basis of normalized population data for IQ. The comparability of these measures, both in terms of their distribution and the characteristics of the measures themselves, is questionable. A similar issue exists regarding comparing the "age at talking" data to the data from the Daniels study, which uses yet another set of measures. As the draft analysis is predicated on combining and comparing these z-scores in various ways to support the risk/benefit analysis, some further discussion regarding how the z-scores are influenced by the distributions and characteristics of the data, and ways in which the analysis and its conclusions might be influenced by these issues, is essential.

Comparison of either z-scores or IQse across endpoints that are qualitatively very different is questionable. The lengthy text on pp. 87-99 describes the mathematical aspects, but not the qualitative aspects of the comparison. It is one thing to make such comparison among psychometric tests (e.g. WISC-R vs. PDI vs. McCarthy scales, etc.), but an entirely different thing to compare the z-score for a psychometric test to a z-score based on observation of developmental stages (which in the case of the Iraqi observations are extremely error-prone). A further consideration is the known error in individual observations of age at talking in the Iraq study population, vs. the error characteristics of the psychometric tests. An expert opinion is needed on whether the similarity in z-score (or IQ-se) for age at talking vs. those from Axelrad et al. and Cohen et al. is meaningful and supports FDA's interpretation.

Description and documentation of modeling approaches remains very unclear (including potential biases/uncertainty):

The addition of Figure IV-2 and associated Table VI-1 go a long way towards more clearly documenting the key assumptions underlying exposure modeling in the FDA model and the implications of those assumptions. However, the presentation of information in this figure and the related table, could be further strengthened if the nature of potential biases and uncertainties in the modeling approaches were discussed. Specifically, do individual biases/uncertainties impact the mean prediction of exposure, or do they more significantly impact high-end percentiles? This additional information helps the reader to understand how biases/uncertainties (associated with specific assumptions) could impact modeling and therefore, could be useful in terms of aiding transparency related to overall uncertainty. In addition, unfortunately, it appears that some of the looping diagrams documenting exposure modeling in detail (present in earlier drafts of the report) have been removed in the latest draft of the report. These more detailed diagrams were critical in allowing the reader to understand the modeling approach and should be added back into the report. EPA comments during the Interagency review had called for these to be clarified and for an actual example of calculations for one or more consumers as illustrated using these looping diagrams, to be provided.

Explanations of the actual steps in the modeling in the Appendix continue to be incomplete and/or very difficult to follow. Three key examples of this are how the short-term CSFII[2] consumption data were expanded (in part, in combination with the NHANES[3] data) to adequately describe the patterns of intake of less frequent consumers; the conversion from variable fish intake to blood Hg concentrations; and the incorporation of the temporal variability in blood Hg levels into hair Hg concentrations. These are critical elements of the exposure modeling that drive the entire modeling process. Notwithstanding the highly questionable aspects of the outcomes modeling, without a valid and clearly comprehensible model of exposure, the model results cannot be interpreted

Methodology used for the quantitative analysis of MeHg effects remains problematic.

[2] Continuing Study of Food Intake by Individuals, a survey conducted on a continuing basis by the U.S. Department of Agriculture

[3] National Health and Nutrition Examination Survey, a continuing study conducted by the U.S. DHHS.

This continues to be a major technical issue. The only major change in the modeling of effects is that, in this second draft, FDA chose to forego separate analyses of risk and benefit for cardiovascular effects and, instead, to treat fish consumption (independent of data on MeHg dose-response or intake) as composite exposure metric. Since FDA attempted to make the case in draft Omega-3 Report that fish had an overall beneficial effect on cardiovascular health, the modeling of this endpoint basically compared the amount of fish intake to the amount of accrued benefit. For neurodevelopmental effects, the FDA draft analysis is essentially unchanged from that in the first draft. The risk portion of the draft analysis continues to rely heavily on the data on age of first talking from the Iraqi poisoning. The justification for this continues to be inadequate and EPA remains seriously concerned about this choice. IQ outcomes are derived from the Faroes, New Zealand and Seychelles, but the translation of the Faroes domain-specific outcomes to estimate IQ measures on the basis of partial correlations remains a major source of uncertainty. The draft analysis continues to use z-scores to integrate disparate neurodevelopmental outcomes (age of first talking and IQ). While there is some statistical rationale for the use of z-scores to integrate across measures, the statistical aspects of the process should be secondary to the conceptual validity of the integration. To put it metaphorically, z-scores can be used to integrate across apple and orange data to produce a measure of variability in fruit salad, but it is first necessary to establish that data on fruit salad are meaningful. An appropriate rationale has not been provided in the current draft for an assessment of neurodevelopmental "fruit salad".

Risks and benefits are modeled from separate databases.

From a structural standpoint, a major problem with the draft analysis is that although it intends to be a risk-benefit analysis, risks and benefits are modeled from separate databases. The rationale for this is that in order to avoid the confounding of risks by benefits and benefits by risks, it is necessary to find data that address risks-only and separate data that address benefits-only. This creates several significant problems.

The first is that it precludes the possibility of addressing an interaction between (e.g.) omega-3s and MeHg. That is, that increasing the exposure to MeHg not only increases the MeHg-specific risk, but also decreases the omega-3-specific benefit. Evidence for such an interaction for both cardiovascular effects and neurodevelopmental effects can be seen in the studies of Virtanen et al. (2005) and Oken et al. (2005, 2008) respectively.

The second problem is that because fish contains both MeHg and omega-3s, it is essentially impossible to find data from fish consumption that dissociate risks and benefits. To deal with the need to identify risks unconfounded by benefits, the FDA draft analysis relies on the MeHg-treated grain exposure from Iraq (Marsh et al.1987). Thus, the draft analysis continues to use the highly uncertain data on age of first talking from Iraq that has been abandoned by the rest of the scientific community for the purposes of risk assessment. Adverse effects of MeHg on age of first talking have not been reported from any other study despite looking for such effects. The obvious conclusion is that, for whatever reason, this is not an appropriate endpoint for fish consumers in the U.S. population. Even if the data on this endpoint are reliable in the context of Iraq, it does not appear that they are relevant to other exposure situations. This may be due to the high-dose nature of the Iraqi exposure, or to the nature of other conditions in Iraq. Furthermore, even within the context of the Iraqi data, themselves, there are insufficient data in the relevant dose-range to

support the level of analysis attempted here. The 2000 NRC report clearly stated that these data are not appropriate for risk assessment for the U.S. population. The attempt to augment these data with data on *verbal comprehension* from the Seychelles further compounds the problem because the Seychelles data are not for the same specific endpoint and not comparable to the Iraqi data. No attempt is made to establish the comparability of these data from a developmental standpoint and it is merely assumed that since both have a verbal focus they can be combined in a "mash-up." Furthermore, the Seychelles data clearly do not meet the *a priori* criterion of presenting risk unconfounded by benefit.

A parallel problem is that it is equally impossible to find benefits data from fish consumption that is unconfounded by risk. The FDA draft analysis uses the Daniels et al. (2004) data for this purpose. The rationale for this choice is that Daniels et al. stated that they detected no significant adverse effect from mercury on neurodevelopment. It seems unlikely, however, that this population in the UK had significant fish consumption but no significant MeHg exposure. There are numerous design, analytical and statistical reasons why an effect of an existing MeHg exposure on the developmental outcomes might not have been detected. This does not mean, however, that the outcomes do not reflect an underlying influence from MeHg exposure.

A clearer discussion of plateaus and thresholds is needed and how they are (or are not) accounted for in the models.

What is the justification for assertion that a plateau of fish benefits must be above the 95th percentile? (e.g. p. 69 second full paragraph.) How does this assumption (no plateau) compare with the decision to incorporate the possibility of a threshold for mercury effects in the model? At a minimum, sensitivity analysis of this assumption should be provided. The failure to include in the model a plateau for fish benefits appears to be contrary to the data (as acknowledged in the document on p. 71; 'Daniels et al. … suggests a plateau'). The predictions of the model regarding benefits are highly impacted by this assumption (see p. 93, where it is stated that "beneficial effects do not exceed the size of one IQ point until consumption exceeds 44.2 grams of fish per day….").

Please note that all of the "what if" scenarios are potentially impacted by the assumption of no plateau for the beneficial effects of fish. It would be useful if an alternative model assuming a plateau effect (which the authors state is supported by the Daniels study) were provided for comparison purposes.

The authors speculate on possible thresholds for the adverse neurodevelopmental effects of mercury, based on results of FDA's model (i.e. pp. 89 and 91). However, page 68 states that threshold assumptions were built into the model itself. EPA's IRIS assessment says that "It is also important to note that no evidence of a threshold arose for methylmercury-related neurotoxicity within the range of exposures in the Faroe Islands study,". Furthermore, Oken et al. 2008 reports: "We observed associations of mercury levels with child cognition at exposure levels substantially lower than in populations previously studied. Our findings suggest that no lower threshold exists for the adverse effects of prenatal mercury exposure."

Alternative risk / benefit analyses were not discussed

The other risk / benefit analyses described in the document need to be described in a more comprehensive manner. Also, several other risk / benefit analyses for methyl mercury exposure in fish have been conducted (e.g., Landrigen et al.) and should be addressed in the text.

Significant limitations are associated with inability to model fish species-specific benefit contribution.

While the model does have the ability to predict differences in risks reflecting different fish species consumption patterns, the opposite side of the equation (benefits) does not. The inability to differentiate fish species in terms of their potential beneficial health impacts is a significant limitation that calls into question the ability to make any risk-benefit tradeoff estimates beyond a simple mean baseline estimate for the U.S. commercial fish consuming population. If it is believed that there is the potential for different fish species to have different health benefits (on a unit consumption basis), then the inability to reflect this factor in the modeling means that there is a high degree of uncertainty of behavior of the FDA's risk-benefit model; this becomes obvious as predictions are generated further from the population mean (or reflecting any form of fish species tradeoff – i.e., the "what-if" scenarios included in the FDA draft analysis). One way to explore the implications of this limitation on the benefits side of the model would be to conduct a sensitivity analysis that reflects potential differences in health benefits linked to different fish species by basing them on the relative content of different species of one or more of the posited health benefit agents (e.g., omega-3 fatty acids, selenium etc). In other words, use the relative levels of one of these agents across the fish species as a proxy for quantitative differences in the magnitude of the public benefit of consumption across species. This would at least allow runs of some risk-benefit tradeoff simulations where fish species are differentiated in terms of benefits based on some plausible metric. While not potentially representative of actual differences in benefits, this approach would provide a means of assessing the potential importance of this factor in impacting risk-benefit tradeoffs (particularly for higher-end modeled percentiles).

Section IV does not contain a separate discussion of results from the Daniels "benefits" model, but only of the results of the combined model.

A separate discussion of the "benefits model," similar to the discussions of the "adverse effects" models, is needed.

There is lack of clarity as to the distinction being made between individual effects and population effects.

This is treated in the draft Risk / Benefit Report as a major issue and is used as a rationale to exclude certain data from the draft analysis. However, it is still not clear what is meant by this distinction. Regardless of whether data are initially expressed as population distributions, or as individual observations, they are ultimately combined into distributions that are sampled probabilistically. Thus, it remains entirely unclear what the distinction actually is between population and individual effects and why it is important.

Results from the models should be presented as the original values (both medians and variance), in addition to the presentation of z-scores (e.g., in tables on p. 125).

Statistical methods exist for modeling individual variability from summarized data

Page 42 of the draft Risk / Benefit Report describes some criteria for limiting the choice of the neurodevelopmental data set. There are statistical methods for modeling individual variability from summarized data for continuous parameters – the authors have not explained why such methods could not be applied. These methods would allow for expanded choices in input data

The discussion of benefits with respect to IQ is inconsistent with the measures modeled in the draft analysis.

It seems inappropriate that much of the discussion of benefits is with respect to "IQ" [and in fact several paragraphs refer to benefits in IQ points (e.g., p. 127), which were not evaluated in the model], rather than using the measures modeled in the draft analysis. In addition, presentation of figures plotting the actual values (i.e., test scores) versus the methylmercury values, along with the modeled data, would allow the reader to better evaluate the model fit (which was not discussed).

Not all neurodevelopmental or cardiovascular endpoints benefit equally from omega-3s or fish consumption.

Neither are all endpoints are equally adversely affected by MeHg. By examining a very limited selection of endpoints, there is no guarantee that the endpoints that are chosen are those that are most susceptible to either benefits or risks. Structuring fish consumption advice on the basis of such a narrow picture of either benefits or risk is likely to result in missing the point of major impact of either.

There is a need for more rigorous sensitivity analysis.

There is great potential benefit of sensitivity analysis aimed at a specific "high profile" aspect of the model (i.e., the potential species-specific differences in health benefits linked to fish consumption). However, there are many other factors in the FDA model, which should be examined through sensitivity analysis, including these: (a) the decision to give several statistical models equal weight in describing mercury concentrations in fish species; and (b) the decision to truncate the hair-to-blood ratio distribution at the 20% upper percentile.

There are also several instances within the FDA model where a specific modeling step has multiple modeling options. In some of these situations, the authors derive a single risk-benefit estimate by giving each alternate modeling option equal weight and including them collectively in a single probabilistic simulation. It is important to recognize that this approach implicitly assigns a confidence level to the underlying models. This is a major step that arguably should not be undertaken without clear rationale for the derivation of the confidence weights, or potentially the use of expert elicitation, conducted in some defensible fashion. In many of these instances, it would be beneficial to step back and examine the impact of these multiple modeling options through both single-factor and multi-factor sensitivity analysis (rather than making assumptions regarding their confidence levels and integrating them into the overall model as a single hybrid-component). The

use of single-factor sensitivity analysis would allow the impact of uncertainty in individual modeling elements to be assessed. The use of multi-factor sensitivity analysis would allow a set of viable risk estimates to be generated (each reflecting a specific set of modeling approaches), which can be interpreted as reflecting a set of uncertainty realizations. While this would not result in specific confidence levels to place on each of the risk results (generated using the multi-factor sensitivity analysis approach) when considered together, they would provide a set of risk estimates that can inform our overall understanding of uncertainty related to the FDA model.

3. Comments on the Neurodevelopmental Assessment

The choice of critical study for the neurodevelopment assessment is critically flawed

The Iraqi exposure episode, from which the neurodevelopmental data used in the assessment were extracted, was an acute, high dose exposure not similar to expected exposure scenarios for the U.S. population (largely consumption over a long period of time of MeHg-contaminated fish). Exposures described in studies of epidemiological cohorts from the Seychelles, Faroe Island and New Zealand are more relevant and appropriate for use in assessing the U.S. population. Further, the Iraqi study is dated, relies on gross measures of effects and the reliability of the dates at which the children reportedly achieved the various gross developmental milestones has been questioned by a number of reviewers. Studies with a high degree of sensitivity were ignored because they did not have individual data. No attempts to get the original data were documented in the draft Risk / Benefit Report.

The continued use of these data by in the draft remains a major gap in logic and scientific justification. The draft Risk / Benefit Report continues to ignore the fact that the Iraqi endpoint of delayed onset of talking (as well as walking) was not reported to be associated with MeHg exposure in any other study. While this endpoint was clearly associated with MeHg exposure in Iraq the appropriateness of this endpoint to other conditions of exposure has not been established. Furthermore, the relative absence of maternal exposure data from Iraq in the exposure level of interest for fish consuming populations in the U.S. has long been recognized. Until this modeling exercise, the Iraqi data had been completely abandoned by the scientific community as a basis for risk assessment for more than a decade. The reasons for this were detailed by several authors and most notably by the NAS/NRC committee on the Toxicological Effects of Methylmercury (NRC 2000). These data are not only of questionable utility from the standpoint of quantitative risk assessment, but their toxicological validity is doubtful for populations at considerably lower levels of exposure. The draft continues to use these data without apparent recognition of these problems.

Criteria for selection of neurodevelopmental effects studies are overly narrow

Some of these are described on pp. 41-42. Several other issues might be considered, including size of cohort, suitability of endpoints, exposure levels, etc. The issues discussed on pp. 66-68 are an indication of how it might have been useful to consider other criteria. This is an example of how the report could have benefited from an initial design-review phase.

Although FDA's exclusion of key studies on the health effects of MeHg is not necessarily a matter of applying explicit exclusionary criteria, FDA continues to exclude the Faroes studies as well as

the New Zealand study (except for some secondary IQ issues). The ostensible reason for this is embedded in the structure of its model for neurodevelopmental risk-benefit. That is, FDA structured its model to assess MeHg effects in the absence of benefit from fish consumption, and benefit in the absence of MeHg effect. The requirement of the model to dissociate risk and benefit resulted in FDA viewing the Iraqi data as the only available data that reflected exclusive MeHg risk since the exposure was not fish-based. Given this view of risk-benefit modeling, FDA's intent was to exclude studies that reflected integrated risk and benefit. Given this structure, however, it is difficult to understand how FDA continued to use the Seychelles developmental data in combination with the Iraqi data

Selection of endpoints used in modeling risk for adverse neurodevelopmental effects is flawed

The endpoints selected (age at first walking and age at first talking) are not the most sensitive measure of adverse neurodevelopmental effects seen in methylmercury exposed populations. More sensitive endpoints of cognitive and sensory function from the Faroe Island study were examined by the NAS (NRC 200) and recommended for use in risk assessment. And FDA says on page 6 of the draft Risk / Benefit Report *"Verbal development is one of many aspects of neurodevelopment. We used verbal development in young children as n indicator of neurodevelopment because we had data on it sufficient to develop; dose-response functions. …is not necessarily the aspect of neurodevelopment that is most sensitive to MeHg."* This is an important caveat. It suggests that this modeling effort was undertaken using this approach because the approach could be supported by the available data. Adverse findings on walking and talking and /or apical measures of IQ are indicative of deficits of higher order process, but they do not obviate that adverse neurological deficits may still be present in the absence of effects on these outcome measures as is the case in numerous studies not selected for risk / benefit analysis in this report.

The critical consideration here is not specifically whether the endpoint on which the risk / benefit analysis is conducted is the most sensitive with respect to MeHg (i.e., to risk alone). Rather, the critical consideration is whether the endpoint is the most sensitive given the net trade-offs of benefit and risk. That is, an endpoint with some degree of MeHg sensitivity may also be one which is very sensitive to the beneficial effect of omega-3s. For such an endpoint, the risk may be cancelled out or even exceeded by the benefits. However another endpoint with moderate sensitivity to MeHg may be one that experiences no omega-3 benefits. The net risk from such an endpoint could be significantly greater than from the most MeHg-sensitive endpoint that also receives significant omega-3 benefit.

Additionally, some discussion of the differences in baseline (i.e. unexposed) 'age at talking' and 'age at walking' for the Iraqi and Seychelles data, and how these differences might impact the results of the modeling (which includes both data sets), should be included.

Explanation is lacking as to how the Iraqi data are combined with the Seychelles data given that the rationale for selecting the Iraqi data is that they represent risk-only while the Seychelles data combine risk and benefit.

This continues to be an apparent major logical disconnect. The current draft still does not provide either a conceptual rationale or a sufficient methodological explanation for the combining of these data. In particular, it remains very difficult to understand how the Seychelles data on age of first

talking can be combined with the Iraqi data given that no such MeHg effect was reported from the Seychelles.

It is not clear as to how the ALSPAC/Daniels data could be used to model beneficial effects of fish in the absence of MeHg. What basis is there for assuming that there was no significant MeHg exposure in this cohort?

Elsewhere in the document, when FDA makes the case that even moderate levels of exposure to MeHg do not counteract the beneficial effects of fish consumption, the Daniels et al. data are used to substantiate this claim given the measurement of MeHg in the cord tissue of a portion of that cohort. Subsequently, FDA uses these same data to model fish consumption benefit in the absence of significant MeHg exposure. Obviously, both conditions cannot be correct. It remains the case that no clear justification for the use of the ALSPAC/Daniels data to model benefit-only has been provided.

Assumption that adverse and beneficial neurodevelopmental effects occur by the same mode of action is not discussed and may not be warranted.

Another major problem with obtaining risk and benefits data from separate sources is that implicit in this approach is the idea that adverse neurodevelopmental effects and neurodevelopmental benefits necessarily map to each other regardless of the specifics of each. Thus, the risks that are modeled are delayed onset of talking from Iraq (notwithstanding the questionable nature of the validity of this endpoint) and verbal performance from the Seychelles and the benefits are IQ score. The draft analysis presents no biological or developmental justification for assuming that these endpoints are, in fact, comparable. In lieu of such a mechanistic justification, the draft analysis resorts to the statistical solution of converting all results to z-scores and comparing outcomes on that basis. Z-scores are useful for comparing across effects when thee is an *a priori* reason to assume that those effects are inherently comparable. That is, one can combine the variance in apples and oranges to obtain a variable measure of fruit salad assuming that fruit salad is, indeed, what you want to measure. In this case, it is not clear what the "fruit salad" of IQ score, delayed onset of talking and verbal performance actually means in terms of benefit and risk.

Descriptions of the procedures and assumptions in the assessment should be modified and expanded to ensure transparency

Several statements in the Executive Summary are misleading; for, example "Methylmercury is neurotoxic at high levels of exposure", fails to acknowledge that it is also neurotoxic at low levels of exposure, as documented in EPA risk assessments, the scientific literature, and the NAS report (NRC 2000). In addition, the ways in which the draft analysis discusses differences in the consequences of exposure to "high" and "low" mercury containing fish are difficult to follow.

Many statements in the draft report should be adequately supported or explained.

Some of these statements are critical to the conclusions of the draft report. For example, on pp. 125-126, the draft states that the model may overstate the adverse effects of methylmercury. The report then goes on to state that use of data from two sources (the Seychelles and the Daniels study)

results in double-counting of both adverse effects of methyl mercury and benefits of fish (i.e., the data are confounded), but no explanation is provided as to why this would likely result in an overstatement of adverse effects (rather than an overstatement of benefits). To improve transparency, a clearer discussion of these tradeoffs should be provided in this and other similar discussions.

The current focus of the draft report on modeling data for developmental milestones is inappropriate in several ways.

The draft Risk / Benefit Report indicates that no reliable distributions/ confidence intervals for walking from the Iraqi data due to imprecision in assessment of this outcome measure following the poisoning. The distributions developed are from Seychelles data set which presents some substantial uncertainties and concerns. These concerns are based on the highly precocial results found on developmental milestones particularly walking in this primarily African derived population sample. It is not clear from the current draft analysis how representative and, thus, predictive these results are for US population. FDA stated that they had not developed population distributions for these developmental milestones from population-based samples in the U.S. population.

Comparison across neurobehavioral measures is highly questionable.

It is unclear whether a z-score calculated using raw "age at talking" data is appropriately compared to a z-score calculated on the basis of normalized population data for IQ. The comparability of these measures, both in terms of their distribution and the characteristics of the measures themselves, is questionable. A similar issue exists regarding comparing the 'age at talking' data to the data from the Daniels study, which uses yet another set of measures. As the draft analysis is predicated on combining and comparing these z-scores in various ways to support the risk/benefit analysis, some further discussion regarding how the z-scores are influenced by the distributions and characteristics of the data, and ways in which the analysis and its conclusions might be influenced by these issues, is essential.

4. The exposure modeling contains significant flaws.

There are substantial unverified modeling assumptions.

Substantial data on intake of seafood by U.S. consumers are available. However, the data analyzed are insufficient in themselves to provide distributions of usual (i.e., individual average daily) intake of different fish species by pregnant women. As a result, the authors have made several significant unverified modeling assumptions.

> Significant among these is the LTSTCR ("long term-to-short term consumer ratio") variable.
>
> The draft is correct in pointing out that short term fish consumption as measured over a 3 day period will not be a precise indicator of long term consumption patterns. The draft also appropriately makes reference to data on consumer's reported consumption over 30 days

from NHANES. However, it is not statistically correct to assume that a parameter such as the LTSTCR can be used to translate an individual's 3 day consumption (from USDA's CSFII study) into his/her long term consumption. The 3-day survey does not provide enough data – regardless of the functional relationship used to transform it – to estimate long term consumption patterns for these individuals. The draft analysis (Figure AA-1) appears to reasonably reflect the 30-day NHANES consumption patterns. Indeed, the draft report indicates that the LTSTCR relationship was selected specifically for this purpose. However, the 3-day individual data records on grams of fish eaten and specific fish species consumed from CSFII are used to project longer term average consumption overall and for each species. Since the USDA data cannot be transformed to indicate usual fish intake for any individual, projected distributions of usual quantities of fish consumed cannot be relied on. The draft Risk / Benefit Report contains additional steps ("Variation in fish species consumed") to try and project distributions of usual intake for individuals who reported consumption of more than one kind of fish within the 3-day survey. However, these further data manipulations do not overcome the limitation of not knowing usual intake for the individuals whose data is then being further modeled. It should also be noted that if these datasets were to be combined (notwithstanding that the methodology for doing this remains unclear) it would first be necessary to establish the compatibility of the data. This remains an important gap

<u>There is no clear explanation and justification for selection of random repetition frequency.</u>

The "repetition frequency" was described as the number of times that a particular meal, by a particular consumer, is repeated even though the repetition is not represented in the short-term CSFII data. The repetition frequency is selected randomly from an assumed distribution. It is not clear how the result can be other than arbitrary. These issues are absolutely critical to the validity of the Bolger and Carrington model and their validity is questionable. That these issues remain a problem is suggested by the fact that even though the draft analysis employs curve-fitting to make the mean daily fish intake from the model agree, more or less, with the NHANES mean intake, the 90th percentile values given in the report (Table V-1) are still 22% different from those reported by Mahaffey et al. (EHP 17:47-53, 2009). It is very likely that this difference is even larger for the higher percentiles of consumption where the risks are most likely to outweigh the benefits. This is also reflected in the model's prediction of blood Hg levels in the population of women of childbearing age (Table V-3). The model predicts blood Hg at the 90th percentile = 2.9 μg/L and at the 95th percentile = 4.3 μg/L. In contrast, Mahaffey et al. (2009) report from the empirical NHANES database that the 89.6th percentile = 3.5 μg/L and the 95.3rd percentile = 5.8 μg/L. These are 17 and 26% differences. In fact, the Mahaffey et al. values do not even fall within the 95% confidence intervals of the draft report's predictions. Furthermore, the differences for higher percentiles are likely to be considerably greater.

Mercury concentrations in Table AA-2 appear to come from "FDA surveillance data, which is not appropriate for reflection of nationwide distribution of fish consumed in the U.S.

While these extensive data likely contain much valuable information for evaluating mercury exposure from seafood, it is not appropriate to assume – in the absence of further evaluation and

substantiation - that a collection of such measurements accurately reflects a nationwide distribution for fish consumed in the U.S.

There are limitations to using 'unit fish' as an exposure measure in the modeling.

While it may be a useful initial modeling construct, the insights that can be obtained from such a model, particularly given the wide range of types of seafood consumed and the differences in nutrient and pollutant concentrations, are relatively limited. The development of models that examine the effects of differing levels of nutrients and pollutants in seafood and different consumption patterns would support modeling with more insight as to health risks/benefits associated with different choices regarding the types fish consumers choose to consume and rates at which they consume commercial seafood. The authors do not review possible differences in risk based on the type of fish consumed. Further, the authors do not review possible differences in risk based on the preparation of the fish (e.g., baked vs. fried). Risk communicators will not be able to use this draft report to help consumers make better decisions regarding the types of fish that consumers should choose to eat, their rates of seafood consumption, or preparation methods.

The issue of representativeness – whether data collected for other purposes can be taken to provide a statistical distribution for the U.S. – should be discussed in the draft FDA report.

Instead of discussing the issue of representativeness, the draft analysis concentrates on conducting a highly complicated analysis of concentration data for each species in which a "battery of 10 distributions was fit to each data set and the four that provided the best fit were used to construct a probability tree." While the draft analysis's attempt to capture uncertainty in potential distributional forms is commendable, in our judgment, such formal and mechanized statistical treatment is computational "overkill" that does not lead to appreciation of the strengths and uncertainties of the database. In particular, if the data are not known to be statistically representative for the U.S., detailed distributional analysis is in the end misleading to the reader.

The discussion of benefits with respect to IQ is inconsistent with the measures modeled in the draft analysis.

It seems inappropriate that much of the discussion of benefits is with respect to "IQ" [and in fact several paragraphs refer to benefits in IQ points (e.g., p. 127), which were not evaluated in the model], rather than using the measures modeled in the draft analysis. In addition, presentation of figures plotting the actual values (i.e., test scores) versus the methylmercury values, along with the modeled data, would allow the reader to better evaluate the model fit (which was not discussed).

The analysis used to predict maternal blood mercury levels as a function of estimated methylmercury ingestion may be flawed.

The hair-blood ratio is not strictly a physiological or toxicokinetic parameter that can be modeled from first principles or that can be generalized from empirical data from a given population. This ratio is a function, not only of the kinetic relationship between steady-state blood Hg and hair Hg, but also (given the kinetic offset between blood Hg concentration and hair Hg concentration) a

function of the frequency of consumption. This was not addressed and continues to be a major source of error and uncertainty

The Sherlock et al. (1984) study, that FDA appears to rely on, appears to be a valuable experimental investigation of methylmercury intake versus blood levels in 20 (presumably healthy) male volunteers. However, the elaborate "uncertainty" evaluation provided in the draft – resulting in developing 120 probability models to describe the Sherlock data seems to miss a central point: to what extent are the experimental data on a "convenience" sample of 20 men representative of the population distribution of blood/intake relationship for methylmercury in the population of American women of child-bearing age – and more specifically on relationships for pregnant women. However, instead of focusing on evaluating the degree to which these data are likely to be applicable to the population of concern, the draft focuses on extensive computer simulations conducted on the assumption that the data are representative. A maternal hair to blood ratio for methylmercury is then needed to complete the exposure conversions. Here the authors had available a substantial and statistically based sample of women of child bearing age from the NHANES database. A population distribution for this variable was taken from the observed hair/blood ratios for the NHANES sample. However, the draft, while correctly recognizing the potential for "noise" (i.e., error related statistical variability) in a study of this nature, deals with this concern by (arbitrarily) deleting the top and bottom twenty percent of the observed distribution of ratios. Where the observed ratios go from values on the order of 0.1 to 5, the truncated distribution extends only from 0.1 to 0.3. While the authors' motivation may be understandable, a pruning of an empirical distribution in this manner does not provide a statistically (or scientifically) valid way to project the actual population distribution.

There is some confusion between methylmercury and total mercury in the analyses.

On p. 56, in the right side of the flow diagram, two boxes shown are "MeHg in Hair" and Neurobehavioral MeHg-Response." The methods employed apparently are attempting to isolate methylmercury exposure alone, rather than total mercury. However, all dose-response relationships are for **total** mercury. This appears to be an inappropriate mixing of exposure metrics. Since there are not dose-response relationships for methylmercury body burdens, it seems that exposure should be expressed in total mercury terms. If the authors wish to use only methylmercury exposure, then it would appear that different dose-response functions would be necessary. At the very least, there should be detailed discussion of these exposure metric issues.

Note also on p. 67, paragraph 3, that the exposure metric in the Seychelles study is total mercury, not methylmercury.

There is lack of agreement between empirical NHANES blood Hg concentration data for women of childbearing age and the modeled distribution of blood Hg.

The mean blood Hg concentration for women of childbearing age appears to agree reasonably well with published NHANES data. However, from the document it appears that the model was adjusted to fit the central tendency of the empirical data. In comparing the upper percentiles of the model predictions to the empirical data, the discrepancies become significant. Mahaffey et al. (2009) report that the 89.6^{th} percentile for blood Hg = 3.5 μg/L and the 95.3^{rd} percentile = 5.8 μg/L. In

contrast, the FDA model predicts the 90th percentile = 2.9 μg/L and the 95th percentile = 4.3 μg/L. It is likely that these discrepancies become progressively larger for higher percentiles. However, since it is clear that the greatest risk of neurodevelopmental and CHD risk reside in the upper percentiles of the distribution of blood Hg in the U.S. It appears that the FDA model under-predicts exposure for the critical fraction of the population. This is remains unaddressed.

5. Comments on the Cardiovascular Assessment

The draft Risk / Benefit Report needs a clear description of the choices that were considered in the modeling

In the 'He analysis' the selection of this data set – rather than others – needs to be logically described. The choice of modeling coronary heart disease in individuals above and below 45 years of age is rather crude. The draft Risk / Benefit Report needs to address, with an explanation, whether or not the data are amenable to analyzing smaller age groups (e.g., 5 or 10 year spans). Addressing age differences is potentially important because coronary heart disease rates and fatal coronary disease risks change over time in the U.S. population. In the modeling based on the He et al data, the dose-response function is forced through the background risk rate at low doses. FDA should be able to show the reader if and how changes in background risks affect the model results. It seems that this would have a big effect on the predicted results. Further, in the FDA's fatal coronary heart disease dose-response model based on He et al. (2004), the authors state that no model uncertainty was included in the analysis. While the authors note that this is consistent with data obtained from epidemiologic meta-regression/analysis, at a minimum this comment should be substantiated with a reference from the peer-review literature and some additional explanation. Given the importance of this draft analysis it was surprising and disappointing that uncertainty analysis was not undertaken. If the FDA has an opportunity to conduct sensitivity analyses, it would substantially increase the quality and potential usefulness of the analyses.

Criteria for study choice and treatment are unclear, but it appears to be flawed.

It is entirely inappropriate to lump all of the studies measuring MeHg and CHD effects together and simply add up the positive and negative studies. These studies were not of comparable quality. For example Guallar et al. (2002) was much larger, better controlled, and appears to be more appropriate to the question at hand than either of the Swedish studies. Yoshizawa et al. (2002) was a well conducted study but cannot be clearly interpreted because of the potential confounding from dentists' exposure to elemental Hg. The draft Omega-3 Report presents arguments as to why the Finnish CHD data are not applicable to the U.S. (e.g., low Se, low omega-3s, high MeHg). While it is not at all clear that any of these are necessarily true (see comments on CHD literature below), the draft Risk / Benefit Report dismisses Guallar et al. even though it does not have these potential problems. Instead the draft Risk / Benefit Report uses the Finnish data that were dismissed in the Omega-3 Report.

The draft report should analyze non-fatal coronary heart disease risks.

It is a serious issue that the draft Risk / Benefit Report does not include a section describing non-fatal coronary heart disease risks. At a minimum, this report should discuss the evidence examining the impact of increased fish consumption on non-fatal coronary heart disease risks. Roughly 20% of all heart attacks in the U.S. are fatal. The influence of commercial fish consumption on the incidence of non-fatal coronary heart disease, roughly 80% of the total, needs to be addressed.

The draft Risk / Benefit Report needs to provide some additional treatment and discussion of the confidence intervals in the 'Carrington model' results.

The reported confidence intervals suggest that there is some likelihood that fish consumption increases coronary heart disease risks. The single sentence in the text ("However, the bulk of the probability distribution is less than zero, so it is more likely than not that increased fish consumption leads to a decrease in CHD death.") was unsatisfying. The lack of discussion regarding this point leads to a broader concern regarding the general classification of the exposure as 'fish.' This is unsatisfying for a number of reasons. The draft Risk / Benefit Report does not review possible differences in the risk based on the type of fish consumed. Further, the authors do not review possible differences in risk based on the preparation of the fish (e.g., baked vs. fried). Individuals cannot use this report for decisions regarding the types of fish they choose to eat or preparation methods. Further, reduction in fatal heart attacks is an important benefit- perhaps the most important in the analyses. This needs additional analysis and discussion in the main text.

It is not clear how the coronary heart disease models treat changes in risk over time

Under the scenario, some women decrease fish consumption while pregnant, then, presumably, return to their pre-pregnant ways consuming more that 12 oz of fish/day. The draft Risk / Benefit Report says that this would lead to a small increase in CHD risk (which is consistent with the authors' assumptions). EPA assumes that this is a one period model and that the model is just for the increased risk during the pregnancy but this is very inconsistent with some of the coronary heart disease models that describe an increased constriction of coronary arteries over time. At a minimum, the modeling and model assumptions need additional explanation.

The draft Risk / Benefit Report incompletely and, at times, inaccurately describes the cardiovascular literature.

In its description of the effects of MeHg on heart disease, the report states that 'adverse associations were seen in two of five study populations but not in the others, including one in the United States. The studies from both of the positive cohorts (Salonen (multiple studies) and Guallar) need to be better described. The Guallar study, in particular, is a multicenter study. FDA continues to incorrectly characterize the influence of "study center" on the findings (ignoring the authors' valid statistical control and sensitivity analysis). Not only was mercury associated with increased risk of heart attacks in the overall study but in nearly all of the centers, most of which were based in Europe. This is significant and needs to be highlighted in the report along with criticisms of the study.

Assertions in the draft Risk / Benefit Report that Finnish populations showing cardiovascular effects of mercury exposure were low in omega-3 fatty acids are not correct. In fact Rissanen et al. (2000) observed a protective effect of omega-3s; they found a 44% reduction in acute coronary events between upper and lower quintiles of serum omega-3 concentration. Obviously there was sufficient omega-3 present in the diet of this cohort to exert a significant positive effect (controlling for MeHg exposure).

The focus on supposed low levels of Se and omega-3s in the fish diet of the cohort in eastern Finland does not have a clear basis in the data. The role of Se in mediating the CHD effects of MeHg is unknown. In Guallar et al. (2002), for example, there is no reason to assume that Se was low given the marine fish consumption by the subjects and the known high Se content of marine fish. Nonetheless, a significant relationship between MeHg and CHD was still seen.

Note on p. 29, par 3 that the discussion of CHD relative to Minamata failed to reference and discuss the study of Tomashiro et al. (1987). This study did, in fact, show a CHD effect of MeHg exposure.

A complete explanation is needed of the studies describing cohorts that saw no association of coronary heart disease with mercury exposure.

There are important details to discuss for two of the three cohorts that saw no association with mercury. If dental workers (who are likely exposed to inorganic mercury) are excluded from the Yoshizawa et al. study, this study reports a positive association of heart attacks with toenail mercury exposures, although not statistically significant. The interpretation of the study by Ahlquist et al. is difficult because they report plasma concentrations, which is a better marker for inorganic mercury exposures rather than methylmercury. This leaves only the Halgren et al. study as being negative. The draft Risk / Benefit Report would be much improved if it provided better descriptions and analyses of each of these studies.

The draft Risk / Benefit Report should use the same outcomes for all coronary heart disease models (e.g., change in deaths could apply to both the He model and the Carrington model).

Delete “averted deaths;”. Should this have been “averted coronary heart disease deaths”?

6. Comments on increasing concentrations of methylmercury in ocean fish

In section II Exposure to Methylmercury in the United States (pp. 13- 15) the draft Risk / Benefit Report poses an important question: “Are Concentrations of Methylmercury Increasing in Commercial Fish?” In answering this question, the draft Risk / Benefit Report contains errors regarding the current scientific understanding of mercury cycling and methylation, thereby resulting in the misleading statement that “limited data suggest that methylmercury concentrations have not increased nor decreased over time.”

It should be emphasized in this section, as is done elsewhere in the draft Risk / Benefit Report, that while most commercial fish eaten in the U.S. are marine fish coming from different

oceans and estuaries around the globe, a small fraction is from domestic freshwater environments. While the draft report's discussion of mercury trends is limited to the open ocean, it should be noted that numerous examples exist in freshwater and coastal marine systems where reductions in atmospheric loads have resulted in lower mercury levels in biota, including fish. For example, it has been found that mercury levels in fish from Swedish lakes have decreased with atmospheric loads (Munthe et al., 2004). Also, a recent article by Monson (2009) performed a trend analysis of the mean mercury concentrations of a standardized length northern pike (55cm) and walleye (40cm) in a set of lakes across Minnesota. In this study, they found that there was a decreasing trend in fish tissue mercury concentrations before the mid-1990s, after which the trend reversed and an increase in concentration was documented. It is unclear what caused the observed trends, but the researcher suggests that the earlier decrease may have been from reductions in local emissions, while later increases may be caused by increased global anthropogenic emissions. Other possible influences that were suggested by Monson include global climate change or changes in sulfate deposition, affecting mercury methylation rates.

With respect to marine fish, however, based on EPA's knowledge, no statistically robust datasets documenting trends in mercury have been developed and, thus, at present, the timing and the magnitude of the response of fish mercury concentrations to open ocean changes in mercury loadings remains unclear. However, modeling studies for different ocean basins show that anthropogenic enrichment of mercury in the water column of the Atlantic Ocean and Mediterranean Sea, for example, is greater than 50 percent (Sunderland and Mason, 2007). In addition, until very recently (Sunderland et al., 2009, in press; Cossa et al., 2009) evidence for an "in ocean" methylation process was not available, but with this recent work a source of methylmercury to marine pelagic food webs has been identified and can be used to put constraints around response times to changes in loads. Rather than inferring from the present lack of data that a "steady state" situation exists, as the draft Risk / Benefit Report does, the situation suggests a need for such data. Indeed, as will be noted, this area is the subject of ongoing investigation.

The draft Risk / Benefit Report focuses its brief consideration of changing methylmercury levels in marine fish on a single study (Kreapiel et al. 2003) that detected no change in Yellowfin tuna tissue concentrations caught off Hawaii once in 1971 and again 27 years later in 1998. Many factors that could have affected mercury levels in this species over this period were not studied, such as changes in growth rates and food structure, and commercial fishing pressures which greatly reduced the average age and size of Yellowfin caught between 1971 and 1998. In addition, modeling studies show that the North Pacific Ocean, where these fish were harvested, will be slower to respond to anthropogenic mercury source enrichment of deposition compared to other ocean regions (Sunderland and Mason, 2007). In its draft report, the FDA presents Kreapiel et al.'s (2003) suggestions that "mercury is converted into methylmercury (the form of mercury in fish) in the deep ocean, with transfer to the upper layer of ocean taking a minimum of 400 years" as established fact. However, Kreapiel et al. (2003) do not actually have any data to support the supposition of deep ocean mercury methylation; rather they infer it from their observed static Yellowfin mercury concentrations. We do not believe their conclusions relevant to mercury and methylmercury sources and cycling are justified by the data of the study. In fact, their hypothesis regarding mercury methylation and cycling is directly contradicted by other reputable scientific findings.

A number of recent studies have collected data directly demonstrating mercury methylation in the marine water column (Cossa et al., 2009, in press; Ekstrom et al., 2006; Kirk et al., 2008; Monperrus et al., 2007a; Monperrus et al., 2007b; St. Louis et al., 2007; Sunderland et al., 2009, in press) indicating that the hypothesis that mercury requires transport to and from sediment and/or benthic food chains in the deep ocean for methylation to occur is most likely incorrect. There are many factors affecting methylation and bioaccumulation that can be expected to result in time delays in the response of various aquatic systems to different atmospheric loads. In a recent paper Sunderland et al. (2009, in press) have reported evidence in marine waters for the importance of particulate organic carbon (POC) transport and remineralization on the production and distribution of methylated mercury species in the water column. Sunderland and Mason (2007), noting wide differences in mercury levels found in different ocean basins, present a box model for mercury cycling in open ocean regions that estimates that the temporal lag between changes in atmospheric deposition and ocean mercury concentrations will vary from decades in most of the Atlantic up to centuries in the bottom waters of the Pacific Ocean. The modeling performed by Sunderland et al. (2009, in press) suggests that basin-wide North Pacific mercury levels may increase over the next several decades if global emissions are only maintained at current levels; acknowledging the present unclear result for fish concentrations, they conclude that "such increases could have serious implications for resulting contaminant burdens in pelagic marine fish if methylated mercury species production mimics total mercury concentration trends."

7. Specific Comments by Page

p.9 par. 3

"The data on exposure presented in this section derives from a national survey of hair and blood levels in the U.S. conducted by the CDC and from FDA's surveillance database on concentration of mercury in commercial fish in the U.S."
Mahaffey et al. (2009) have shown significant regional differences in MeHg intake. Summary NHANES data can therefore be misleading since (consistent with the overall averaging approach in this draft analysis) there is considerable population variability in MeHg exposure that is not reflected in the summary NHANES data.

p.11 par. 2

"Because NHANES is designed to provide a nationally representative picture of exposure in the U.S.…. it dies not lend itself to regional analysis ". This does not recognize the recent Mahaffey et al. (2009) publication that does provide regional data. Ignoring these regional differences would result in underestimating risk to significant fraction of the population.

par. 3

NHANES national focus would appear to reduce its ability in any assessment of risk for localized situations…However, the limitations do not significantly affect the utility of NHANES in a nationally representative assessment of risk… Modeling that FDA has performed …closely track body levels as reported by NHANES."
It is not surprising that modeling efforts aimed at reproducing the average U.S. consumption would closely track the national average NHANES data. That is not the relevant point. The

relevant point is that modeling the average consumption and then using these results in the development of fish consumption advice will mean that the resulting fish consumption advice will not necessarily be protective of those whose consumption does not coincide with the national average.

p. 13 3rd bullet

The value of 0.35 ppm Hg for canned albacore tuna is the value reported in the FDA database for 2002-4. However, Burger and Gochfeld (Env. Res. 96:239-44, 2004) reported a mean of 0.41 ppm with 25% > 0.5 ppm.

5th bullet

There are several fish that have a significant regional or ethnic-based consumption (e.g. bluefish in the northeast; carp among Asian consumers; halibut in the northwest) that should be included among the mid-range species. This is another example of the information lost in focusing this draft analysis on the national average consumption.

p.21 par.5

Hibbeln et al (2007) did not, in fact, have a measure of MeHg exposure.

p.20 par. 3

Statement re overt neurologic abnormalities observed at levels 100 time average body levels in the U.S. is misleading. Note that overt symptoms have been reported in U.S. cases wherein the blood Hg levels were measured as low as 38 µg/ l.

p.22 par.1 and ff.

It is critical to understand that the relationship between fish consumption and neurodevelopmental outcomes (risk or benefit) in these studies is specific to the mean fish consumption pattern (and the mean consumer) in each specific cohort. Without the ability to statistically link specific effects to specific characteristics of the fish diet (e.g., omega-3 intake, MeHg exposure) there is no basis for generalizing to other populations and especially to specific upper percentiles of other populations.

Par. 3

"*In the New Zealand study…the overall net effect from eating fish was not measured.*" This is not entirely correct. Phase 2 of the study included a high-fish/low-Hg group. Results were not reported as stratified by fish consumption, but these data can be recovered from the study reports and reanalyzed to determine the effect of fish consumption.

p. 23 Table IIIA

For the Iraq study (Marsh et al., 1987), the 'finding' category states that the significant adverse association was found between neurological exam results and milestones and prenatal exposure. It does not state that authors statistical analyses evaluated only the overall neurological exam results, not the milestones used for the FDA endpoint (i.e. age at talking and age at walking).

Budtz-Jorgensen et al. 2007. First bullet, sentence 2: the statement re "fish contribution to net effect independent of methylmercury" is unclear, and seems to be proposing an interpretation of the analysis that the paper's authors did not provide. Sentence should be deleted.

p. 24

The findings for the Oken et al., 2005, study are converted into gains and decrements of IQ points per weekly fish serving, without referencing the FDA authors' assumptions of a specific mercury level for the fish. The actual 'gains and decrements' would be very dependant on how much mercury was in the specific fish consumed.

For Oken, 2008, the findings section states that 'average benefits were lower still when mothers ate no fish during pregnancy'; this would seem misleading in that it would be impossible to 'benefit' from eating fish if no fish are eaten.

First bullet phrasing, especially "reduction in those improvements" is a very confusing statement of the results, and differs from the characterization by the study's authors. Oken et al. plainly state that higher mercury levels were associated with poorer cognitive test scores." The sentence should be revised to: "Maternal fish consumption was associated with improvements on the tests while mercury body burdens were associated with reductions in test scores."

Oken et al. 2008, bullets 2-4. Specify that these findings are for the WRAVMA total score. No such analysis was provided for the PPVT score. From the information presented in Figure 1, it appears that the findings on these points might be different for the PPVT (i.e. each test shows a similar effect for mercury, but the fish benefit is reduced for the PPVT).

p. 25 Table IIIA-2

Lederman et al. 2008. This study does not belong in this table – it is not a study in which exposure to methylmercury was not measured (as stated in the table caption). This study belongs in Table IIIA-1. (Note that the summary of findings in the table states: "the study reported that mercury was associated with lower scores." Given that Lederman is a study of a U.S. population consuming a mainstream fish diet that has measures of both neurodevelopment and MeHg exposure, it would appear to be the ideal population for this draft analysis.

p. 26

The discussion of postnatal exposure on p. 26 fails to note that the children studied in the Faroe and Seychelles Islands also had prenatal exposure, making it difficult to evaluate the effects of postnatal fish consumption.

The Daniels study is used to model 'fish only' benefits, for the risk / benefit comparison that is the main focus of this document; however, on p. 26 it is stated that 'it is necessary to assume that increases in postnatal fish consumption in this study population were accompanied by increases in methyl mercury exposure.' If it is true that increases in fish consumption are confounded by increases in methyl mercury exposure, it would seem necessary to make some attempt to include this confounding variable in the benefits model; there is no indication that such an adjustment was made.

p. 27 par. 2

The lower concentration of MeHg in maternal milk compared to maternal blood is not a transport issue. Rather, it results from the fact that milk is derived from plasma. Plasma is enriched in inorganic Hg compared to the erythrocytes and relatively depleted in MeHg.

p. 30 par 3

"Although these studies [studies assessing the relationship between fish consumption *per se* and CHD effect] *did not measure MeHg levels in individuals…it is reasonable to assume that the fish contained MeHg."* This is a reasonable assumption, but the fact that the fish contained <u>some</u> unknown level of MeHg is beside the point. The critical information is the level of MeHg compared to the omega-3s. Without knowing this, these studies provide no useful information for the construction of fish consumption advisories. It is also pertinent to ask why, in the case of these CHD studies, FDA is assuming (in all likelihood, correctly) that for CHD modeling, populations consuming a significant amount of fish must also have had a commensurate level of MeHg exposure, while in the case of neurodevelopmental modeling, the draft report assumes that the Daniels et al. (2004) cohort had a significant consumption of fish with no significant MeHg exposure.

p. 41 1st bullet

"methylmercury effect not confounded." The draft report apparently interprets this to mean studies where the methylmercury exposure did not come from fish. Another option for selecting studies that would provide much more data would be to use studies that statistically controlled for fish benefits (i.e. Budtz-Jorgensen et al. 2007, Lederman et al. 2008, and Oken et al. 2008).

2nd bullet

"… we had to model some aspects of neurodevelopment that we could assume to be reasonable indicators of at least part of the MeHg adverse effect on neurodevelopment as a whole." This raises several critical concerns. Given the multifaceted nature of neurodevelopment, what basis is there for the assumption that there are individual endpoints

that are, in fact, qualitatively and quantitatively indicative of the whole? FDA provides no justification that the endpoints that are chosen are the most sensitive to MeHg. Second, what is the basis for assuming that every endpoint reflects the same balance of MeHg and nutrients (e.g., omega-3s)? Endpoints that are not particularly responsive to omega-3s may, nonetheless, be sensitive to MeHg. Such endpoints may be the overall most vulnerable even if they are not the most sensitive to MeHg in isolation.

p. 42 par. 1

"Individual variability cannot be modeled from summaries of data because summaries presume a distribution… that precludes the possibility of modeling individual variability…" This is unclear. What is meant here by individual variability? How can there be individual variability in fetal neurodevelopment? Each individual has a single outcome value for each outcome parameter. Furthermore, doesn't the existence of a summary "*distribution"* imply variability that has been modeled?

par. 2
"We were especially reluctant to use statistical summaries that had been subject to a log (dose) transformation because the impact of the transformation on the secondary modeling results is difficult to determine." This is opaque.

p. 43 par. 2

"in four of the five studies that looked at MeHg and CHD, data… were obtained through methodologies tat make comparison of exposures… difficult. These methodologies involved measuring MeHg levels in toenail clippings… Without the ability to make such comparisons, it is not possible to know the MeHg levels in the study participants as revealed by established biomarkers, e.g., whole blood and hair." This is incorrect. Ohno et al. (2008) provide a direct quantitative relationship between toenail Hg and hair Hg.

p. 46 consumption, implications

"*…there is some indication the fish consumption in women of childbearing ate may have decreased since the CSFII survey was conducted."* Mahaffey et al. (2009) have recently shown using NHANES data, that fish consumption in this group has not declined although MeHg intake has decreased.

short-to-long term frequency, implications

"Persons who consume seafood very rarely (less than once per month) are not well characterized." Why is <1/mo considered the point at which uncertainty about consumption patterns of infrequent consumers becomes significant? The gap between 1/3 days and 1/30 days is quite significant and logically would become a problem at about 1/wk.

% of consumers eating fish over an entire year, assumptions

"As part of the long-term correction, an adjustment is made to account for the fact that the number of fish consumers is increased as the length of the survey period increases." However, as the length of the survey increases, not only does the number of consumers change, but the type of consumption (species portion size) may also change. Contrary to the "Implications" note, this occurs over a much shorter period of time than a whole year. The issue at hand is not the % of consumers over the course of the year, but the range of consumption patterns. This issue becomes clear if one thinks in terms of species and types of fish consumed rather than merely thinking of fish as a "package."

p. 47 long-term species consumption patterns, assumptions

"…market share data can be used to reasonably determine varied consumption." Market share data can only determine the varied species consumption averaged across the population. This should not be confused with population variability in consumption. These data can provide no information on the extent to which some consumers consume a greater proportion of high MeHg fish than the average.

p. 48 serving size adjustment, assumptions

"…current per capita consumption is more accurately measured by market share disappearance…" This is only true if one is dealing with estimating the average serving size. Market share data cannot provide data on population variability in portion size.

p. 49 par. 6

"*We used these* [NMFS] *data to help estimate the types of fish consumed over a year."* Since the NMFS data do not relate directly to individual consumption, they cannot reflect population variability in consumption. At best they can only help in establishing the average fish diet.

"NMFS market share data were also used to adjust portion sizes to reflect current levels of consumption." Again, since NMFS data do not relate to consumption patterns of individuals, they can only estimate the average portion size and cannot provide information on variability in portion size.

p. 50 using 30-day survey

The repetition ratio is not well explained here or in Appendix A. Contrary to this text, math underlying this ratio is not presented in Appendix A.

Using the three-day survey and the NMFS market share data, why would it be assumed that the sum of individuals consuming a specific type of fish with a specific portion size would be described by a random repetition ratio for that combination of species and portion size? Rather, it would seem that different species and portion sizes would be associated with

specific repetition ratios (i.e., these are not independent parameters). Canned tuna is a prime example of this.

p. 52 par. 4

Give rationale for body weight $^{0.44}$.

p. 58 blood-hair relationship, assumptions

"That the relationship between blood and hair Hg is presumed to have the same proportion at all doses (i.e., linear). The hair-blood relationship may be linear across doses (i.e., across blood concentrations), but it is not linear across dose *ratios.* It varies by consumption frequency. This is not taken into account.

p. 60

The assumption that "the results have not been substantially confounded by methylmercury" is contrary to available data; for example, both Budtz-Jorgensen (2007) and Oken (2008) were able to separate beneficial effects of fish from adverse effects of methyl mercury in fish-consuming populations. Since no threshold has been demonstrated for adverse effects of pre-natal exposure to methyl mercury on neurological function, a more appropriate assumption would be that there was some confounding, which would need to be addressed as part of the risk / benefit model. This has been addressed in a number of recent papers, most recently Ginsberg and Toal, (2009).

The implications column states in points (3) and (4) that problems regarding confounding and dosimetry (i.e. the assumption that all fish have similar effects) is taken into account in the confidence intervals. As the confidence intervals provided (for some parameters, only) are presumably for the model itself, not the data (or assumptions) used in the model, some additional explanation as to how the calculated confidence intervals reflect the assumptions would be useful.

choice of indicator for neurobehavioral benefits, assumptions (3)

"minimal confounding by MeHg." If data are from fish consuming populations, then how can MeHg confounding be considered minimal since it is inseparable (except statistically) from fish consumption?

choice of indicator for neurobehavioral benefits, implications

"*the results form the modeling are generally consistent with studies… involving other aspects of neurodevelopment measured at later ages in life."* This is confusing. If these endpoints at 15 and 18 months were considered consistent with results form testing in later life, then why not use test results from later life (e.g., 4 years, 15 years) wherein tests are more sensitive to both benefits and risks?

Shape of the relationship between fish consumption and neurobehavioral benefits, implications (4)

"The fact that the dosimetry treats all forms of fish equally means that the model does not differentiate between species with respect to neurobehavioral benefits. The implication is that net effects could vary from diet to diet, within the range of confidence intervals."
The nature of these confidence intervals has not been clearly presented and it is hard to see how the confidence intervals were constructed, which could lead some to view them as arbitrarily constructed.

p. 62

Note several problems with numbering of points in Table IV-2

Point 15

1. Assumption (a) *"age of talking is a useful indicator of neurodevelopment."*
It is completely illogical to compare benefits on verbal comprehension to age of first talking (and certainly to first walking). These are not necessarily the same endpoints and they come from completely different studies of very different levels of exposure. Furthermore, and critically important, this endpoint was only observed in the high-dose exposure from Iraq. It was not reported in any other study and the dose-response for this endpoint is highly uncertain in the dose range of interest.

2. Assumption (c) regarding use of 'age of talking' from the Iraqi data does not address the lack of precision in the Iraqi data (i.e., the birth dates of the children were unknown, the age of talking was retrospectively determined during interviews with the mothers), or that the article specified that 'age of talking' less than 24 months was considered normal for this population. In addition, another measure from the MCDI, verbal production [included in the appendix tables on p. 146] would seem to be more analogous than verbal comprehension [the endpoint used to calculate 'benefits'] to age of talking. No discussion was provided as to why comprehension was selected instead of production.

3. Implication for (c) does not address the issues noted above. In addition, based on the graphs on p. 187, the Seychelles Island data show no apparent effect on 'age of talking' up to the highest exposures shown. The lack of effect in the Seychelles data, when combined with the differences in recorded 'age of talking' [virtually all the Seychelles children appear to be talking by 15 months of age, whereas even the least exposed Iraqi children appear to be rarely talking by 14 months of age] also raise concerns regarding the appropriateness of using these combined data sets to represent the adverse neurological effects of prenatal methyl mercury exposure.

p. 63

Point 15 (second one)

1. Assumptions:

a. Assumption (a) regarding confounding appears to be faulty, and has been discussed above with respect to the benefits analysis.
b. It is unclear why a single linear model was deemed appropriate for the benefits analysis but for the adverse effects analysis multiple functions (linear and nonlinear, threshold and non-threshold) were used. In addition, the method used to generate the Z-score slopes is unclear (see also comments on appendices).
c. Although the assumption states that several types of analyses were used, as best we could determine only the analysis based on the Iraqi/Seychelles age of talking data was used in the combined risk/benefit analysis.

2. Implications
   a. The comparative analysis referred to in (a) (and again on p. 67) was not provided in the document. We have previously provided comments on issues regarding confounding and differences in the underlying data for the Iraqi and Seychelles studies.
   b. The response in (b) does not address the validity of extrapolating down to effects in days when the underlying data have a precision only of 6 months, or the validity of extrapolating such high dose data to the low dose region of concern.

p. 64

Point 16

1. The Assumptions point understates the extent to which the comparison (i.e. with Axelrad, 2007, Cohen, 2005, etc.) used to support the current analysis is dependent on the use of Z-scores, which in turn are dependant upon the standard deviations (s.d.) chosen to calculate the Z-scores for the Iraqi data-based model.
2. The 'implications' makes several points which appear to be inaccurate (see comments above about the apparent differences in baseline 'age at talking' in the Iraqi and Seychelles data). In addition, authors of the Iraqi paper noted that rural Iraqi mothers did not talk to their infants very much, and defined as 'normal' age of first talking up to 24 months (p. 1021, Marsh et al., 1987). This would not seem consistent with the Seychelles data (see above). The appropriateness of using the SD from the Seychelles data for this population is questionable.

p. 66 par. 1

"*It is highly uncertain whether separation of such highly correlated variables can be done.*" That is, in fact, what has been done for omega-3s and MeHg in two recent papers (Strain et al, 2008, and Budtz-Jorgensen et al., 2007). Use of these analyses would eliminate the need to use the largely irrelevant Iraqi data and combine it with the benefit data from elsewhere.

par. 4

It is difficult to imagine how this source of uncertainty can be in any way acceptable. One would assume that this would preclude the use of these data in the draft analysis.

p. 67 par.1

Were data requested from the Faroes Islands studies? It seems no more of a stretch to compare milestone data between Seychelles and Faroes than to consider age of talking and results of verbal comprehension tests to be equivalent.

"*If we had used the milestone data from the Faroe Islands, we would have had to do so in lieu of the Iraq data.* This does not seem to be much of a rationale, particularly given the deficiencies of the Iraq data.

par. 3

Data on the age of talking in the Seychelles have not been published and no citation is given. There is no way to evaluate those data. However, given the lack of any reported test deficits in children below the age of six (at least from the longitudinal study), it is difficult to imagine statistically significant developmental delays as a function of maternal MeHg.

2nd bullet

Presumably, no association was found between MeHg exposure in the Seychelles and age at first talking since this was never reported despite explicit investigation. Therefore, there is doubt as to the validity of this endpoint for lower (than Iraq) levels of exposure – especially since there does not appear to be an effect at any such dose. Thus, using the Seychelles data introduces a bias toward a decreased association. Furthermore, the only possible rationale for the use of the Iraqi data is that it is not confounded by fish-related benefits. Given that, why were the Iraqi data combined with data from the high fish-consuming population in the Seychelles?

If (as stated) the addition of the Seychelles data does not make a significant difference in the outcome variable (presumably delayed developmental milestones), then why confuse the already complex nature of this analysis?

p. 68

Several of the above assumptions are addressed in the discussion at the top of p. 68. Again, the actual procedures used to select and validate the model are not clearly described, and many of the comments above are pertinent to the discussion in this paragraph as well.

par. 1

"The second assumption is that MeHg might have a threshold effect." Given the fact that the Iraqi data reflects what is largely a very high-dose exposure, and given the lack of data in the range of the presumed threshold, models assuming a threshold are likely to be uncertain to the point of being entirely uninformative.

par. 2

"A fourth assumption is that ages of first talking and walking are useful measures for neurological health. The salient point here is not whether these are necessarily useful

measures of neurological health, but whether they are sensitive indicators of MeHg effect. Since they have not been reported to be associated with MeHg exposure except in Iraq, it is hard to see how they can be used in the analysis of relatively low-dose exposure in the U.S.

p. 69 par. 3

Replace each instance of "methylmercury" with "mercury." Axelrad et al. used dose-response data from the Seychelles, Faroes and New Zealand studies. Each of these studies measured total mercury, not methylmercury

p. 70 par. 2

This text is difficult to interpret. A clear statement of the assumption and its implications is needed. We suggest: "The analyses conducted by Axelrad et al. and Cohen et al. used data on the relationship between mercury body burdens and neurodevelopmental test performance that did not control for the beneficial effects of fish. Recent studies (including reanalysis of Faroe Islands data (Budtz-Jorgensen et al. 2007) as well as Lederman et al. 2008 and Oken et al. 2008) demonstrate a stronger effect of mercury when the model accounts for benefits of fish. Therefore, the dose-response estimates of Axelrad et al. and Cohen et al. will underestimate the adverse effects of mercury."

par. 3

The majority of MeHg in the Faroes Is. was from the consumption of pilot whale rather than from fish. The nutritional profiles of pilot whales and fish are not the same… so there was less opportunity for confounding by nutrients in fish…" This doesn't make sense. The Faroese are high fish consumers. The fact that their MeHg exposure comes largely from whales does not alter their intake of fish nutrients.

In New Zealand, the high consumption of shark says nothing about consumption of other fish or fish nutrients. There are no data from New Zealand to support any conclusions about fish nutrient exposure. For the Seychelles, the reasoning regarding the confounding by fish benefits is backwards. It is assumed that there were few fish benefits because the IQ dose-response was adverse. This presumes that if there were fish benefits, then there would be no discernable negative IQ effect, but that is precisely the concept being tested. One cannot assume the hypothesized outcome in order to test that outcome

Basis for statement "suggesting that fish confounding was not substantial" in the Seychelles is unclear. Statement appears to be entirely speculative. Final sentence minimizes potential impact of fish benefits on mercury coefficient estimation without support. A clear and objective discussion of this issue is needed; or better would be to consider a model in which mercury coefficients controlled for fish benefits are used.

p. 71

At the top of the page, authors state that their "benefits of fish" model was not confounded by the methyl mercury present in fish consumed by the subjects of the Daniels study. No basis given for this statement.

par. 3

What is justification for assertion that plateau of fish benefits must be above the 95th percentile? How does this assumption (no plateau) compare with the decision to incorporate the possibility of a threshold for mercury effects in the model? At a minimum, sensitivity analysis of this assumption should be provided.

p. 72

We note that, contrary to the assertion at the top of p. 72, the Iraqi data used in the 'adverse effects of methyl mercury' model are episodic (i.e. the results of a single poisoning episode), not long term.

p. 74

Why is the PPVT in Oken et al. 2008 not considered to represent early age verbal?

p. 78 par. 4

The distinction between the CHD pooled analysis model and the CHD meta-analysis model is unclear. Since both the meta-analysis model and the pooled analysis models combine data, it is not clear how the pooled analysis model can any better reflect individual population-specific risk factors than the meta-analysis model.

p. 85 par. 2

Mahaffey et al. (2009) (*Environ. Health Persp*. 117:47-53) report a 90th percentile fish intake for women of childbearing age of 25.2 g/day compared to 32.3 g/day in Table V-1. This suggests that the model becomes uncertain and unrealistic for higher percentiles. This is particularly important since it is in the higher percentiles where we would specifically expect to see adverse effects.

p. 87 Table V-3

Table V-3 needs to make clear (in title and/or notes) that these estimates are for hair mercury due only to intake of dietary methylmercury from fish, and do not represent total mercury body burden.

Mahaffey et al. (2009) report from analysis of the NHANES data that the 89.6th percentile of blood Hg for women of childbearing age is 3.5 μg/L and the 95.3rd percentile is 5.8 μg/L. In Table V-3, the comparable percentiles are 2.9 and 4.3. Furthermore the confidence

intervals reported in the table do not include the Mahaffey et al. values. Thus, these models do not accurately reproduce either fish consumption or MeHg exposure in the upper percentiles of consumption and exposure where the adverse effects are most likely to occur. This raises serious questions about the accuracy of the model predictions.

pp. 87-99

Comparison of Z-scores (or IQ-se) across endpoints that are qualitatively very different is questionable. The lengthy text here describes the mathematical aspects, but not the qualitative aspects of the comparison. It is one thing to make such comparison among psychometric tests (e.g. WISC-R vs. PDI vs. McCarthy scales, etc.), but an entirely different matter to compare the Z-score for a psychometric test to a Z-score based on (extremely error-prone, in the Iraqi case) observation of developmental stages. A further consideration is the known error in individual observations of age at talking in the Iraq study population, vs. the error characteristics of the psychometric tests. An expert opinion is needed on whether the similarity in Z-score (or IQ-se) for age at talking vs. those from Axelrad et al. and Cohen et al. is meaningful and supports FDA's interpretation.

p. 89 par. 1

This analysis applied to a MeHg developmental endpoint that has only been found once in a high-dose exposure to a low-dose exposure and found that MeHg in this low dose range has little or no effect relative to that endpoint. All that this shows (at most, assuming that the analysis, itself, is valid) is that this is not a low-dose-sensitive effect. It says nothing in a larger sense about the effect of MeHg from fish consumption in the U.S

p. 90

The footnote for Table V-4 indicates that the FDA percentile values for mercury are lower than NHANES (some NHANES values are provided in Appendix B, pp. 170-171, not referenced here), and provides a possible explanation for this difference. To improve understanding of how the conclusions of the risk-benefit analysis are impacted by the differences in exposure estimation, it would be helpful if comparable estimates based on the NHANES data were provided.

p. 91

We note that the 'change in IQse' based on the age of first walking data (Table V-5,) is twice that based on the age at first talking data (Table V-4). This difference reinforces our concern regarding the sensitivity of the model outcomes to the specific neurological endpoint selected to represent the 'adverse effect of methyl mercury'. Presumably if the age of first walking endpoint were used instead of the age of first talking, the amount of fish consumed in order to achieve a net benefit would be larger (again depending on the methyl mercury content of the fish). Further analyses exploring this issue are needed.

par. 2

Note again that this analysis purports to demonstrate a threshold for a frank effect, grossly measured.

p. 92

The document does not indicate whether the Axelrad and Cohen analyses were adjusted to account for the 'benefits of fish'. If not, then a direct comparison of the results in Table V-6 to the authors' analysis would be inappropriately confounded. In addition, it would be useful to include variance estimates in TableV-6, rather than just central estimates (especially as two different analyses are listed in the Table title).

p. 93 par2

Assuming that this is, in fact, a valid analysis, these results suggest that IQ is not a sensitive indicator of the beneficial effects of fish – as opposed to e.g., VRM, or the individual domain tests in the Faroes.

p. 94 par. 1

"*The Net Effect on Fetal Neurodevelopment from Commercial Fish In order to estimate the net effect on fetal neurodevelopment from maternal consumption of commercial fish we developed this model by combining the results from age of talking in Iraq and Seychelles (representing MeHg) with early age verbal comprehension results…*" This is misleading. In describing this as the "net effect" on fetal neurodevelopment, it implies that together, IQ and age at first talking constitute the entirety of neurodevelopmental endpoints affected by either beneficial aspects of fish or MeHg risk. This is not the case. Foe example, the Faroes study has shown effects on memory, attention, motor skills and coordination.

p. 96

The statement at the top of the page "one-tenth of one percent of the population is likely to experience an adverse effect and that most of the remainder of the population is likely to experience a beneficial effect…" appears to be somewhat biased – an equally accurate statement would be that the results indicate that half of the population will experience little or no benefit or have a negative net effect. This is in spite of the fact that a linear response (with no plateau) is assumed regarding beneficial effects of fish.

p. 100

Minor math errors in last column of table. Row 1, 0.225 should be 0.255. Row 2, 0.0105 should be 0.015

p. 108 par. 2

Which dose response function was used for all four sub-populations?

p. 115 par. 2

It seems illogical to conduct an analysis using 1989-91 (i.e., 19 year old) data given that one stated concern is that fish consumption has decreased in response to the current advisory. There are no bases given for an assumption that fish consumption, both qualitatively and quantitatively, has not significantly changed during that time?

p. 116 par 4 and ff.

"Because short term surveys are better at monitoring consumption patterns for frequent consumers than for infrequent consumers, the LTSTCR in serving frequency was reduced for frequent fish consumers." How were frequent and infrequent consumers identified in the data considering that the data only address three days of consumption? These adjustments remain unclear and seem arbitrary.

p. 118

"This distribution was derived from the NHANES survey by calculating the fraction of total fish consumption in the fish category with the highest number of eating occasions for the 403 adult women who consumed fish on four or more occasions." Our reviewers could not follow this despite repeated efforts.

Fig. AA-2

What is "Major Category Frequency?"

p. 119 par. 1

Our reviewers could not follow this.

par. 2

Our reviewers could not follow this.

"Individual variation in species consumption and overall frequency of consumption were assumed to be independent." This does not make intuitive sense. People who eat fish most frequently are people who tend to eat the same fish repeatedly – most specifically, tuna.

par. 3

"A concentration factor was applied to serving sizes to reflect water loss during food preparation." It isn't clear that this is appropriate because it isn't clear if people report portion size based on purchase weight or cooked weight. However, since fish is bought by weight and weighed in the store, and not weighed post-cooking, it would seem more likely that the adjustment results in a bias toward decreased fish consumption.

par. 4

"A correction factor of 1.15 was applied to portion sizes form the CSFII survey to that total intake matched per capita estimates from NMFS…" This is an arbitrary adjustment since it is not known whether misreporting of portion size is uniform across species and types. In particular, canned tuna is less liable to portion size bias since can weight is uniform.

p. 127 par. 4

This does not explain how individual variability in blood Hg based on the frequency of MeHg intake was accounted for. This is an important consideration since most consumers are not at steady-state with respect to MeHg in blood and detailed information on the frequency of consumption by species was not available. Furthermore, it is not clear how body weight was accounted for since CSFII does not supply body weight. And why is it advantageous that corrected values (using $BW^{0.44}$) have no correlation with body weight?

p. 131 1st bullet

Blood and hair Hg levels follow different kinetics. The ratio of blood Hg to hair Hg, therefore, is very sensitive to the frequency of consumption. Levels <1 ppb reflect infrequent consumption and >1 ppb reflects more frequency consumption. Thus, this approach will significantly skew estimates of the relationship between these two markers.

7th bullet

"Regardless of the explanation, because actual pharmacokinetic variation is almost certainly narrower than the apparent distribution, the distribution was truncated with uncertainty ranges of 20 percent." If arbitrary bounds are imposed on the distribution, what is the point of using an empirical distribution?

p. 133

The description of the methylmercury and neurological endpoints dose response function on p. 133 was unclear. Reference to Appendix C (discussed later) did not provide sufficient clarification.

p. 134

These sentences are incorrect: *"Axelrad et al. (2007) used a Bayesian analysis to generate an estimate of a single slope of -0.153 and confidence intervals based on the standard error of the mean that ranged from -0.047 to -0.259. That estimate is employed in our analysis as a normal distribution, per Axelrad, with a mean of -0.153 and standard deviation of 0.064."* The slope and confidence interval provided here are from a sensitivity analysis that used maximum likelihood estimation rather than Bayesian analysis. The primary analysis estimated a slope of -0.180, and a confidence interval of -0.009 to -0.378. These values are

clearly stated in the abstract and in the Results and Discussion sections of the article, as well as Table 5. If FDA used the incorrect values in the subsequent analysis, extensive corrections of tables and text will be needed.

On p. 134 it states that 3 months were subtracted from the reported milestone ages in the Iraqi data due to reporting imprecision. It is stated that this had "little impact on the model fitness the other parameter estimates", but no comparison of the results of these two models is provided. We note that the statement that "onset of walking and talking in Iraq was recorded in sixth month increments" (p. 133) does not appear to be consistent with the data recorded in Marsh et al. (1987), which includes ages reported at 2 month intervals (see pp. 1018-19). As the provided justification for this data manipulation appears to be invalid, it is critical that more explanation be provided, along with the results of the comparable model runs. We previously noted the apparent differences in initial age of talking based on the figures provided in Appendix C. No information is provided as to whether the ages in these figures (pp. 187-188) are based on adjusted or unadjusted data.

pp. 134- 135

The description of the Cohen et al. work is inadequate. Use of test results from all domains (not just cognitive tests) and the role of expert judgment in applying weights to different domains, etc. were critical elements of this study

p. 135

The NAS report (NRC 2000) does not use the phrase "biologically implausible" in its discussion of the log transform used in the Faroe Islands study. The NAS report does discuss biological plausibility of the log transform, but FDA overstates the conclusions. See NAS pages 294, 297 and 315

On p. 135 the document says that study authors state that the developmental delay seen in two Iraqi children was equivalent to 18 standard deviations; this finding is then used to support some of their model choices. We note, however, that the standard deviation they cite is based on the Seychelles data, not the Iraqi data, and, thus, its use in this manner is not valid.

p. 136

It is not clear why the Poland study is represented here, but other studies such as Oken are not.

p. 138

We found the comparison of the neurodevelopmental dose response functions, starting on p. 138, to be confusing as written. Since this section provides a key rationale supporting the appropriateness of the models selected for this analysis, transparency in this section is essential. In the absence of a clear explanation of what was done, along with a clear

rationale, it is very difficult to judge the validity of the analysis as performed. Some, but not all specifics follow.

1st bullet

The text needs to clarify whether the functions used by FDA (Figures AA-11 and AA-12) come from the original authors (e.g. Myers 2003), or from other subsequent analyses of these data by Cohen et al. and Axelrad et al. The figures imply that these functions did not come from data as reported by the original authors; it is not clear why the figures reference Axelrad and Cohen as the source for dose response functions from the New Zealand, Faroes and Seychelles studies.

The meaning of this statement is unclear: "only this group of dose-response functions is consistent to what was observed with the Iraqi study where there are data to anchor the high dose estimates." Bullet 1 appears to make the circular point that as the only high dose data being used are the Iraqi data, they are the only data that are consistent with the effects observed at high doses in the Iraqi study.

2nd bullet

FDA is apparently assuming that the Iraqi data are correct and realistic; if other studies are not consistent with Iraq, they are presumed unrealistic. This presumption appears to skew the subsequent findings. This seems to be stating that because the Cohen 2005b study and the New Zealand data yield slopes different from the Iraqi data, they are unreliable; although the Poland data were also 'higher' than Iraq, they were closer, and therefore not 'necessarily unrealistic'. This appears to be a somewhat biased analysis, based on the assumption that the Iraqi data are more reliable than the other studies. But one would expect the slope from the Iraqi data to be lower as this is a frank effect.

3rd bullet

It is unclear how the Seychelles coefficient described here differs from that in bullet 1. Are two different functions being used to represent the same thing? Why?

Also the 3rd bullet seems to acknowledge possible confounding of the Seychelles data by a 'beneficial effect of fish consumption that equals or exceeds the negative effect of methyl mercury', but does not then explain why it is appropriate to include these data without adjustment as the lower part of the dose-response curve in the 'adverse effects of methyl mercury' model.

last par.

Unclear what "the three dose-response functions considered in this draft analysis" are. Also, how does this statement relate to Figure AA-13, which shows results for two dose-response functions? Further, it is not clear whether/how this discussion relates to the inputs

used in the modeling of neurodevelopmental effects. The discussion concludes without a clear statement of mercury dose-response model inputs.

p. 139 Figure AA-11.

It is very difficult to trace back the various functions listed in the legend to previous text/tables. Greater transparency in showing the inputs, and how they were derived, is necessary. Do all the values in the figures come from the three studies listed in Table AA-6? This is not at all clear. It s not clear why multiple functions are shown for Seychelles, New Zealand, and Faroes – if all need to be included, provide a table that explains to the reader why these studies each have different estimates. Then, only two functions are shown in Figure AA-13 – what happened to all the others?

At the top of the page, there is discussion of uncertainty in the various models and reference to a comparison of confidence intervals in Fig. AA-12, however the referenced figure includes no confidence intervals. As an understanding of the reliability and comparability of the models depends on an understanding of the uncertainty of the estimates provided, this crucial missing information needs to be included in the document.

p. 140

We note that the obvious differences in model predictions on p. 140 are not reflected in the discussion in the body of the paper. More discussion as to why these different predictions are not considered equally valid would be appropriate.

p. 144 – 165 (Appendix A)

Many problems of clarity were noted in the discussion of the model for Fish consumption and neurodevelopmental endpoints, based on the paper by Daniels et al. (2004). Again, a few examples are provided.

It is unclear why MDCI comprehension (what words does the child understand) was selected for comparison, as it would appear that MDCI production (i.e. what words does the child speak) would be a more appropriate comparison/parallel to age of first talking. Although no clear explanation was provided for the various parameters included in Tables AA-7 through AA-10, it is clear that there is a large difference in the sloe estimates for MCDI comprehension and MCDI production, especially with respect to their association with Cord Mercury in Tables AA-7 and AA-8. No discussion of this was provided, but it would appear that this critical decision could substantially impact the results of the draft analysis. A sensitivity analysis comparing the impact of the selected parameter on the model results would seem to be appropriate. We also note that no estimates of the confidence interval around the slopes were provided for any of these variables. This information is important in assessing the reliability of the model outcomes.

The calculation of the ‘Maximum Z-score’, presented as an equation on p. 145, is not clear. No information regarding the meaning of ‘variable range’ is provided, and no values are

provided for this parameter. Tables AA-11 and AA-12 include calculated 'maximum z-score contributions from maternal fish consumption' for each variable, but no explanation is provided as to how these numbers should be interpreted as a 'contribution', or where the values came from (and no confidence limits are provided). Similarly, the source of the numbers in Table AA-13 'Summary of Z-score slopes …' is unclear, nor is the definition of a 'Z-score slope' provided (and, again, no confidence limits). Since these are apparently the values used in the model simulation for this endpoint, some clearer explanation of what these numbers are and where they came from would seem to be necessary. In addition, some explanation should be provided as to why a distribution of Z-scores slopes was used in the benefits model, while there is no reference to such a derived value being used in the 'adverse effects' model.

No explanation is provided regarding how the 'benefits' and 'adverse effects' models were combined.

The meaning of the apparent confidence intervals shown on Fig. AA-14 are unclear, given that the legend indicates that only the median response is displayed and that there is a hidden negative component at the low end of the population distribution. A clearer explanation is needed here, along with some indication of the confidence limits for this combined analysis.

pp. 168 – 179 (Appendix B)

No further explanation was provided here regarding the endpoints selected for use in the model (see above comments).

The results tables in this appendix (e.g., AB-4 and AB-5) would be much easier to follow if the values were expressed in decimal notation rather than in exponential notation.

Questions regarding whether Z-scores were appropriately used here were already discussed above, in particular whether the use of the Seychelles SD for the Iraqi data results in appropriate comparisons.

The interpretation of the last column in tables AB-8 and AB-9 was unclear. Since the whole table is labeled as Z-score change, no explanation is provided for what the 'Net Verbal delta-Z represents.

We also note that tables for different scenarios include different columns, but no explanation is provided for the differences. For example, Cohen (2005b) IQ is included in Table AB-12, but not Tables AB-8 and 9.

We also note that data presented for all intervention scenarios are based on consumption only of low-mercury fish, which is probably not a realistic scenario. Scenarios based on consumption of average-level or high-level mercury fish should also be provided, at least for comparison purposes.

pp. 172 – 176

It is very difficult to follow the presentation in Tables AB-4 through AB-12, making it difficult to determine whether the analysis is reasonable. Column headings need to be more descriptive.

p. 172 Table AB-4

First column heading needs to be more specific: percentiles in what distribution? Were all z-score changes were calculated, then arrayed in a distribution; or are these percentiles of fish intake, or mercury body burden?

Does column two represent estimate of mercury effect without considering fish benefits, and column three the effect of mercury and fish benefits combined? It would be helpful if the column headings made this clear.

For the note to the table – is this referring to the third column, rather than the second? Are only median estimates of the dose-response functions used?

p. 172 Table AB-5

Results shown can not be verified without clarification of column 1 – percentiles of what distribution? Elsewhere, the report has described the Cohen and Axelrad IQ dose-response functions as being very similar, so it is not clear why they have different results in this table.

p. 174 Table AB-8

Why is Cohen included as a column in previous tables, and not here? It is difficult to trace the relationship of this table to previous tables. Do values shown in the Axelrad IQ column incorporate fish benefits? Labeling of this column needs to be clarified. How are the values in the last column (net verbal delta-z) derived (e.g. do they incorporate Carrington estimates, or Axelrad estimates?), and what do they represent?

p. 175 Table AB-12

Again the table format is changed (different columns), making it difficult to understand what is being changed, and what are the results, as one goes from one analysis to another.

pp. 180 – 191 Appendix C

We find the initial paragraph under sources to be somewhat misleading. Although it states that concern for exposure to mercury is based primarily on two poisoning episodes in Japan and Iraq, in fact there is a large literature on adverse effects of exposure to mercury and methyl mercury (some of which are referenced in the second paragraph).

The description of the construction of the comparative modeling (p. 181) does not provide a clear explanation of exactly how the models used in this draft analysis were constructed and selected, or why they are more appropriate than other types of models. In addition, no explanation is provided as to why the more complex model was used for modeling the Iraqi/Seychelles age of talking data but a simpler model was used for modeling the Daniels et al. (2004) Verbal Comprehension data. Since these models form the heart of the risk/benefits analysis, more explanation as to why they were constructed in this way, and the results of these selections on the outcome of the analysis, would seem to be needed.

Goodness of fit is addressed on p. 183, but no information regarding the actual fit of the model is provided. In addition, the document states that fit was based on the number of subjects, resulting in more weight for the Seychelles data in the low dose regions. No explanation is provided as to how the apparent differences in age of talking baseline values were addressed in the model (other than the arbitrary 3-month adjustment to the Iraqi values, mentioned previously), given the large influence of the Seychelles data in the low dose region.

The last paragraph on p. 185 discusses decisions that would need to be made based on the differences in study populations between Iraq and Seychelles data. No information is provided as to what decisions were made in the analysis discussed in the document, and how these issues were addressed.

As noted previously, the figures on p. 187-8 [AC-2 and AC-3] readily demonstrate the large differences in baseline values for the Iraqi and Seychelles populations with respect to the variables modeled for this draft analysis. In addition, neither variable seems to show any effect of increasing exposure in the Seychelles data. The document provides no discussion of this, in particular with respect to why it was considered appropriate to combine these two very different data sets.

Tables AC-2 and AC-3 (pp. 190-191) are difficult to understand and need more explanation in the text.

Addendum1: Critique of FDA Draft Omega-3 Report (Summary of Published Research on the Beneficial Effects of Fish Consumption and Omega-3 Fatty Acids for certain Neurodevelopmental and Cardiovascular Endpoints)

Note that EPA used the services of a contractor in the preparation of these comments on the FDA Omega-3 Report. This was done under contract EP09H000646. The curriculum vitae of the reviewer, Alan Stern, can be found at the end of the critique.

Section A: Cardiovascular

General Comments

The draft Omega-3 Report is completely revamped from that which appeared as and appendix to the earlier FDA draft analysis. Whereas the focus of the literature review of cardiovascular effects in the first draft was on both the risks and benefits of fish consumption given the presence of MeHg and omega-3 fatty acids, the focus of the review in the current draft is entirely on the cardiovascular benefits conferred by fish, fish oil and omega-3s. The clear intent of this revised focus is to facilitate the approach in FDA's modeling exercise of treating fish as a "package" that implicitly integrates risk and benefit without having to dissociate the two. The reasoning being that if it can be demonstrated in the literature review that fish and its fish oil/omega-3s confer overall cardiovascular benefit, then the modeling needs only to address the extent to which increasing fish consumption increases cardiovascular benefit. Furthermore, if fish *per* se is assumed to confer benefit, it can be assumed that an average fish diet will also confer benefit. There are several fundamental problems with the scope and inherent bias of the draft Omega-3 Report. The primary problem is that the great majority of studies on the relationship between fish consumption and cardiovascular effects have been conducted to test the hypothesis that fish confers overall cardiovascular benefit. To produce a result that is consistent with this hypothesis, these studies need only identify an overall average benefit in the population under study. As long as the majority of the study population eating the average fish diet experiences cardiovascular benefit, the overall effect appears as benefit. A fraction of the population eating a fish diet that diverges from the average diet could experience overall increased risk rather than benefit. However, this would be reflected in the population-based outcome as a decrease in the overall benefit rather than as an increase in risk across the population. Studies demonstrating overall benefit to a given population based on its implicit average fish diet cannot necessarily be assumed to apply equally to other populations or fractions of populations that consume a fish diet that significantly diverges from the implicit average diet. In the U.S., there is a range of market availability of different species of fish, and dietary alternatives that allow consumers to vary their intake of fish by both species and frequency. There is, therefore, considerable population variability in fish diets. Thus, all fish diets cannot be considered equivalent with respect to the components that may confer either risk or benefit and not all consumers can be considered to derive equal benefit even if they all consume the same amount of total fish with the same frequency. This is particularly true if, as appears to be the case, the balance of cardiovascular risk or benefit depends on the ratio of MeHg and omega-3s in the overall fish diet. Studies that do not dissociate the contributions of each component cannot be used to predict the balance of risk and benefit that would occur for diets whose MeHg/omega-3 ratio diverges from that in the implicit average diet. If the ultimate goal of fish consumption advisories is to provide guidance that is protective of the average consumer eating the average diet,

then information obtained from the review of such literature may be used to support that goal. Providing that the average diet reflected in these studies can be assumed to apply to the average diet of the recipients of the advisory information, such guidance may protect the average consumer. However, if the goal of fish consumption advisories is to provide guidance that will be protective for all consumers, then conclusions drawn from studies that focus exclusively on testing a hypothesis of overall benefit cannot provide guidance that will be protective of those consumers whose diet diverges from the implicitly assumed average diet. Furthermore, such studies cannot be used to set bounds on the extent of divergence from the presumed average fish diet that will still confer benefit. Since the characteristics of the fish diets in these studies with respect to relative frequency of different species of fish is generally not known, the conclusions about the benefit conferred by an overall fish "package" have an implicit but unknown window of dietary variability within which benefits rather than risks can be anticipated. Therefore, even if the goal of fish consumption advisories based on such studies is to protect the average consumer consuming the average fish diet, the dietary leeway inherent in such advice is unknown.

The second fundamental problem with the draft Omega-3 Report is that in its focus on demonstrating overall benefit from fish consumption, it gives short shrift to those few (but important) studies that attempt to dissociate risk and benefit based on separately examining the effect of MeHg and omega-3s (e.g., the series of studies from eastern Finland, Guallar et al., Yoshizawa et al.). In part, this is the result of lumping these high quality studies together with several other studies that purport to investigate both risk and benefit. Little attempt is made to determine either the extent of their commonality in research hypothesis and design, or the relative quality of these studies. By viewing all of these risk-benefit studies as a whole, the draft Omega-3 Report reaches the conclusion that there is overall inconsistency in the studies that have investigated cardiovascular risk from fish consumption. This entire category of studies is then dismissed as being uninformative. The multi-center Guallar et al. study is singled out for additional and biased critique. Despite explicit statements to the contrary and a sensitivity analysis by the authors that specifically addresses the point, the review incorrectly claims that the findings of the Guallar et al. study are unduly influenced by the contribution of the two centers with the highest MeHg exposure. The Guallar et al. study is further marginalized by confining the modeling exercise to fatal CHD, thus eliminating the need to address this Guallar et al. study since that study focuses on non-fatal CHD.

A third important problem with the draft Omega-3 Report is that much of the available literature focuses on controlled studies of fish oil or omega-3 supplementation. The draft Omega-3 Report makes a reasonable case that fish oil/omega-3s, by themselves, generally confer benefit with little evidence of adverse effect. The implication of the considerable literature that is presented to support this conclusion is that since fish oil/omega-3 confers cardiovascular benefit without risk, and since fish contains these substances, fish, likewise, confers benefit without risk. It is easy to lose sight (especially with the short shrift given to the studies that provide evidence of cardiovascular risk from fish) of the fact that fish contains substances other than these largely beneficial nutrients. Studies of fish oil/omega-3 supplementation in the absence of significant fish consumption can provide no information about risks associated with actual fish consumption. If the modeling conducted to support fish consumption advice proceeds from the assumption that fish confers only cardiovascular benefit and not cardiovascular risk, then studies of fish oil/omega-3 intake are relevant for supporting this position and maximizing the benefit. However, if the

possibility of a risk-benefit balance is accepted, and if, in addition the possibility of interaction between benefits and risks exists, then, at best, these studies can only partially elucidate the benefit side of overall cardiovascular risk-benefit.

In addition, the focus on overall benefit tends to shift the focus away from individual endpoints that either do not experience benefit or experience relatively less benefit from fish, fish oil or omega-3s. This becomes important if one approaches these studies from the standpoint of their implications for risk-benefit rather than benefit-only. From a risk-benefit perspective, endpoints that experience little benefit are more likely to experience risk, and endpoints that experience no benefit are subject to naked risk. The effect of fish consumption on these endpoints, however, will be missed if the focus is on those endpoints that do experience benefit. An important example of this is non-fatal versus fatal CHD. The review makes a strong case that consumption of fish and/or intake of fish oil/omega-3s reduces the risk of fatal CHD. However, the review notes that the evidence for a reduction in the risk of non-fatal CHD is weaker or absent. It may be that to some extent these endpoints have different etiologies and do not exist on the same continuum of risk. Thus, it is ironic that the modeling exercise specifically rejects the Guallar et al. study because it focuses non-fatal CHD rather than fatal CHD. That is certainly not to say that fatal CHD is not an appropriate endpoint for the reduction of risk. Rather, the concern is that a fish consumption strategy that focuses only on the potential reduction in the risk of fatal CHD may inadvertently increase the risk of non-fatal CHD by ignoring the risk of non-fatal CHD that can occur from the same diet that confers benefit-only for fatal CHD.

<u>Specific Comments</u>

p. 9 par. 4 “

In contrast to the large, recognized body of science reporting the cardiovascular health benefits of seafood consumption, a few studies have suggested that MeHg in fish may increase cardiovascular risk.” This is a biased and judgmental editorial statement. It implies that the studies that find risk from MeHg in fish are in contradiction to the studies that do not find risk. As explained in my general comments, there are several reasons, including the specific hypothesis being tested and the existence of risk to segments of the population, why some studies show risk and others do not.

“Thus, the cardiovascular benefits associated with fish consumption in these studies are the net result of beneficial components in fish together with possible risk from harmful components such as MeHg.” This statement is only true regarding the mean outcome of each study. The fact that for some combinations of fish type, MeHg concentration and intake rate, there is a net benefit does not mean that all such combinations confer benefit.

p. 23 par. 4

“Men with higher DHA levels… in the Euramic case-control study or with higher DHA blood levels in a Finnish cohort study had lower risk of myocardial infarction (Guallar et al. 2002; Rissanen et al. 2000). In these two studies, the reduced risk associated with n3 LC PUFA was attenuated by high Hg content in fish. However, SACN emphasized that the

two studies… still represent a positive association between fish consumption and CHD". This is a mischaracterization of these studies. While they did show that n-3s (omega-3s) have an underlying beneficial effect they also show that MeHg has an adverse effect and that the overall balance of risk and benefit depends on the balance of the intakes. With respect to the Finnish studies, this is particularly clear when considering the entire set of related studies rather than from consideration of Rissanen et al. in isolation. The IOM report draws a similar conclusion about the balance between risk and benefit regarding the Finnish studies. From the standpoint of the balance of risk and benefit, fish consumption is neither inherently beneficial, nor inherently adverse.

p. 35 par. 4 and ff.

The summary of the IOM report is a partial quote that is biased toward suggesting that fish consumption is necessarily associated with a decreased risk of cardiovascular risk While recognizing that observational studies have produce "mixed results," the IOM report (pg. 138) concludes that, "Overall, the data considered suggest an increased risk of myocardial infarction among men with higher hair mercury levels." The draft Omega-3 Report does not reflect this conclusion.

p. 52 par. 1 (bullets)

The conclusions of the what-if scenarios from Cohen and Gray based on the Carrington and Bolger model of 2004 do not address how MeHg intake would change with changes in fish intake. That is, how would the *type* of fish consumed change if overall fish consumption either increased or decreased? This is a key question if it is recognized that fish is not a homogeneous food with respect to either risk or benefit.

par. 3

The threshold estimate for omega-3 intake and reduced CHD death in the analysis of Mozaffarian and Rimm was based largely on clinical trials and studies that did not involve MeHg exposure or (to a lesser extent) had an unknown MeHg exposure. Therefore, this estimate addresses benefit-only in the absence of risk.

p. 55 par. 1

"*…most, but not all, reviews and reports have concluded that the studies [of cardiovascular risk from MeHg in fish] are inconsistent and that there is not sufficient evidence to conclude that MeHg in fish increases cardiovascular risk."* The consistency or inconsistency of these studies is an irrelevant point since most of these studies are not comparable in terms of populations, design or quality.

par. 3

The fact that the elevated risk in the Virtanen et al. study was mostly seen in the highest tertile of MeHg exposure is neither here, nor there, with respect to the relevance of this risk

to U.S. consumers. The relevance or lack of relevance depends on the relative magnitudes of exposure. In that regard, the statement that the mean exposure in upper third of the Finnish cohort was 10 ppm in hair is not found in the Virtanen et al. study. In the earlier FDA draft, this was identified as an estimate by the authors of the FDA review document. In this version, it is unattributed (with the implication that this information was in the original Virtanen study) and is not substantiated. Furthermore, it seems highly unlikely to be correct given that the reported mean and standard deviation are 1.9 and 1.9 ppm in hair respectively.

p. 56 par. 2

"Although formal statistical tests [in Guallar et al. 2002] found no effect no effect modification (interaction between Hg level and center, this does not rule out the possibility that results from Malaga or from both Spanish centers may have been influential in the reported positive association between toenail Hg and heart attack risk and led to an incorrect association for the full study population." This statement completely ignores the statement of Guallar et al. in the paper that removal of the Spanish centers from the analysis did not affect the outcomes. Furthermore, this statement mischaracterizes the statistical analysis of Guallar et al. and confuses the issue. Guallar et al. did not investigate effect modification/interaction on the model predictions with respect to the influence of "center." Rather, they conducted a straightforward test of the significance of "center" as an independent variable in the prediction of the outcome (i.e., the odds-ratio).

par. 3

"Additionally, SACN (2004) and Mozaffarian and Rimm (2006) noted that, even in the Eastern Finland cohort and Euramic study… the net effect of fish consumption was still beneficial: greater mercury exposure lessened the benefit associated with consumption of fish of n-3 LC PUFA but did not increase overall risk."

As detailed below by Stern (2007) in a direct response to this statement, the statement is simply incorrect:

> "I can find no basis of support for that statement. Rather, in their Table 4, Guallar et al. show that without adjustment for DHA (i.e., when both Hg and DHA exposure are considered together), the odds-ratio for MI increases dramatically in the highest quintile of Hg exposure (1.0–1.47, p for trend = 0.01). While without adjustment for Hg (again considering Hg and DHA exposure together), the odds-ratio for MI does not differ across the quintiles of DHA (1.0-0.8, p for trend = 0.23). That is, DHA is not protective against the increased risk of MI due to Hg exposure. A downward trend in the odds-ratio with increasing DHA is only seen *after* controlling for Hg exposure. That is, only when the effect of Hg is held constant across levels of DHA exposure is an *underlying* protective effect of DHA seen. Likewise, in Tables 2 and 3 of Virtanen et al., the authors report that without adjustment for DHA+DPA (i.e., considering both the PUFAs and Hg), the relative risk for acute coronary events in men in the upper third of Hg exposure is increased 55–60% compared to men in the lowest third of exposure. Furthermore, the relative risk of acute coronary events with increasing DHA+DPA does not extend below 1.0 when hair Hg concentration exceeds 2.03 ppm (the upper third of the distribution of hair Hg concentration in the study population). These results show that in this population, when Hg intake was

moderately elevated, Hg increased the risk of an acute coronary event and that risk was not offset by the PUFA intake."

It is interesting that this response was not discussed or referenced in the draft Omega-3 Report.

p.57 par. 2 and ff.

Meta analyses are useful in examining a given prediction of risk if the studies that are being combined are essentially comparable with respect to the parameter of interest. For the purposes of this discussion, the parameter of interest is MeHg exposure. These meta-analyses do not, in fact, address MeHg directly and the inherent assumption is that, on average, all studies have equivalent MeHg exposures. Thus, the outcome is relative to the mean MeHg intake among the individual studies and the outliers with higher MeHg intake are ignored. If everyone ate the mean fish diet, these analyses might, indeed, suggest that fish consumption gave greater benefit than risk, but clearly we are concerned with the upper percentiles of MeHg exposure rather than the mean exposure.

p. 58 par. 2

The same argument that is applied here to suggest that exposure misclassification attenuates the estimate of the true magnitude of the beneficial effect of omega-3s, also applies to estimates of the association between MeHg exposure and risk. Budtz-Jorgensen et al. have shown that even with MeHg biomarkers, there is still substantial error in effect estimates for MeHg associated outcomes.

p. 61 par. 3 and ff.

These summaries do not provide information about what type of intake (i.e., fish oil, fish consumption?) the decreased risk was associated with or how much intake.

Section B: Neurodevelopmental

General Comments

Although this section has similarities to Section A (review of cardiovascular benefit) in largely focusing on benefit, it differs somewhat in devoting more discussion to the potential adverse effects of MeHg. While the draft Omega-3 Report makes a reasonable case for the beneficial effects of omega-3s on some specific endpoints of neurodevelopment, it has some important gaps and biases with respect to generalizing and applying such conclusions to the modeling of risk-benefit from fish consumption and ultimately, to the structuring of fish consumption advisories.

As with the review of the cardiovascular benefits literature, this section tends not to do a good job of distinguishing among endpoints. In the case of the cardiovascular review, however, the number of potential endpoints was small. Here, the overall category of neurodevelopment is much broader than cardiovascular health (and more specifically for the review, of CHD). Neurodevelopment encompasses several broad categories including cognition, attention, behavior, and motor skills. Within each of these broad categories, there are specific domains. Each of these categories and

domains has a greater or lesser degree of independence from the others with respect to etiology as well as with respect to its response to both omega-3s and MeHg. The draft Omega-3 Report reveals that not all neurodevelopmental endpoints show benefit from omega-3s or risk from MeHg, and not all beneficial endpoints have comparable levels of benefit in response to omega-3s. Thus, an endpoint that experiences little or no benefit from omega-3 intake may still experience deficits from MeHg. Since, at least in some cases, omega-3s appear to provide protection against the adverse effects of relatively low levels of MeHg exposure, one endpoint may show an overall benefit in response to a given ratio of omega-3 and MeHg intake, while another endpoint may only experience significant risk from the same exposure. Thus, generalizations about the balance of risk and benefit based on a compartmentalized view of benefit may result in unintended harm even if the correct balance of risk and benefit is indentified for the endpoint under consideration.

A particular failing of draft Omega-3 Report is that the case for benefit derives largely from studies of omega-3 benefit in the absence of MeHg exposure or without specific quantitative information about MeHg exposure. There is essentially no explicit consideration of the range of neurodevelopmental effects associated with differing ratios of omega-3s and MeHg. Even when neurodevelopmental benefits are evaluated relative to fish consumption (rather than omega-3s *per se*), there is generally little or no information on the level of MeHg exposure or omega-3 intake experienced by consumers of that fish diet and generally no consideration of the *variability* in MeHg exposure due to dietary variable within and across populations. This is particularly important given the potential variability across endpoints in both omega-3 benefit and MeHg risk. Given the lack of characterization about the components and distribution of components of fish diets in specific studies, conclusions about the balance of risk and benefit derived from such studies should not be generalized to other populations consuming different diets. This also applies to sub-populations within a given study consuming fish diets that differ significantly from the average fish diet in the study population. If such sub-populations comprise a small to moderate fraction of the total, adverse neurodevelopment effects they experience will not be apparent in the mean population outcome and may register as reduced overall benefit. However, the modeling approach used by FDA in the draft Risk / Benefit report document assumes that there is a single average fish diet and bases what-if scenarios on that diet. These scenarios not only assume that all consumers eat the same diet, but also assume that changes in fish consumption in response to advisories will only result in changes in the amount of consumption and not in the type of consumption. This approach ignores the likely importance of omega-3/MeHg ratios.

Essentially all of the studies considered in the draft Omega-3 Report looked at omega-3s in isolation, at fish consumption in isolation (i.e., in the presence of variable, but unknown levels of MeHg exposure, or (in a small minority of cases) fish consumption and MeHg together. None of the studies considered in the review specifically examined omega-3 intake and MeHg exposure together (with the exception of a brief mention at the very end of the draft Omega-3 Report of the studies cited below, without consideration of their implications). This is a critical point, because if risk-benefit information is to be generalized as fish consumption advice, it has to address the specific factors that affect risk and benefit. Advice based on fish *per se* will encompass a wide range of omega-3 and MeHg intakes and thus, cannot address a specific risk-benefit balance. In order to do this, it is necessary to evaluate data that relate the joint, but independent effects of omega-3s and MeHg in the same population. Given this, it is highly significant that the draft Omega-3 Report did not include recent papers that examined a range of neurodevelopmental

endpoints as dependent variables in regression (or structural equation) models that included MeHg and omega-3 intake (i.e., Strain et al. - Neurotoxicology. 2008 Sep;29(5):776-82.; Budtz-Jorgensen et al. -Environ Health Perspect. 2007 Mar;115(3):323-7). These models allow the measurement of the effect of varying MeHg exposure while holding omega-3 exposure constant, and of the effect of varying omega-3 exposure while holding MeHg constant. Since both of these papers have been available for some time, it is not clear why the draft Omega-3 Report (at least the current draft of FDA assessment) did not assess the implications of these papers.

Specific Comments

p. 91 par. 4 and ff.

There is no mention at all of MeHg in this rather long summary of the IOM's report entitled "Seafood Choices: *Balancing Benefits and Risk.*" Although the FDA review specifically states its goal of summarizing the IOM's evaluation of health benefits (and not risk). This seems like an intentional and unnecessary bias given that the IOM report addresses MeHg risk at length.

p. 98

In summarizing IOM's consumption guidance, the draft Omega-3 Report fails to note that the guidance that IOM endorses is the current joint FDA-EPA advice that the FDA modeling is seeking to overturn.

p. 119 par. 3

"Some studies suggest that differences in cognitive function might be transient, a possibility noted also for effects on visual function." This is an important caveat to beneficial neurodevelopmental effects in general, since underlying MeHg effects that are masked by a transient developmental benefit may not, themselves, be transient.

"The authors suggested that global tests such as the Girffith, Bayley or Brunet-Lezine scales might be too broad, including functions that are not sensitive to DHA status." This is also an important caveat for neurodevelopmental testing of omega-3s in general. That is, that not all functionalities that may be sensitive to MeHg will necessarily benefit or receive protection against MeHg from omega 3s.

p. 126 and ff.

The discussion here of the benefits of postnatal supplementation with omega-3s make a reasonable case that at least some postnatal endpoints experience benefit from postnatal supplementation with omega-3s. It is notable, however, that this discussion does not address benefits and risk-benefit considerations for developmental effects of pre-natal omega-3 from maternal intake. Thus, while the conclusions of this section regarding post-natal supplementation may be valid, they are misleading with respect to the possibility of

risk / benefit (as opposed to merely benefit) since there is evidence of pre-natal, but little evidence for post-natal developmental effects of low-moderate doses of MeHg.

p. 134 what-if scenarios

These summaries of IQ gain and loss are based on a model that addresses nationwide shifts. However, this does *not* imply that the same balance of risk and benefit would apply to individual consumers since the scenarios imply that increasing fish consumption would affect only the amount consumed and not the type consumed It is likely, however, that increasing fish consumers for some consumers would mean disproportionate increase in MeHg compared to omega-3s.

p. 136par. 4

In the studies available to Cohen et al., the dose-response for MeHg is presumably confounded by benefit from the concurrent exposure to omega-3s. Put another way, it is likely that the overall effect on neurodevelopment of a particular baseline diet for the what-if scenarios is determined by the MeHg/omega-3 ratio. Even if the effect of the baseline diet on neurodevelopment were known, it is likely that the MeHg/omega-3 ratio would be different under different what-if scenarios. Without knowing the effect of different ratios on neurodevelopment it would not be possible to make a correct estimate of the resulting change in risk (or IQ points).

p. 139 par. 1

Although the Helland et al. (2003) study is key to the overall conclusions in the review regarding the role of omega-3s on neurodevelopment, it is important to note that that study is of fish oil rather than *fish*. Therefore, this study can address only benefits and not risk. It is therefore, not applicable to the mixed benefits and risk situation that applies with fish consumption.

p. 154 par.3

[In the Daniels et al. (2004) study] "Cord tissue mercury levels were available for a subset of 1,054 women. Mercury levels were positively associated with frequency of fish consumption but there was no association between mercury levels and the developmental test scores." This statement is not correct and was not specifically claimed by the authors. Rather, Daniels et al. used fish consumption rather than Hg cord tissue concentration as the measure of combined risk and benefit. Based on the data they present, however, self-reported fish intake was not associated with cord tissue Hg concentration. Categories of increasing fish consumption frequency did not result in increasing cord tissue Hg concentration.

p. 155 par. 2

Contrary to the statement in the review, Hibbeln et al. could not, in fact, adjust their findings for MeHg intake. Hibbeln et al. used the same self-reported fish consumption data as Daniels et al. (see previous comment). Their attempt to derive MeHg intake from self-reported fish consumption was subject to the same lack of demonstrable association between these two metrics. This was pointed out in the literature by Stern and Rice (Lancet. 2007 Jul 21; 370(9583):217-8;). The draft Omega-3 Report should have noted this critique.

par. 3

The draft Omega-3 Report should mention that in the Oken et al. (2005) study, there was an interaction between fish consumption and hair Hg such that those with high hair Hg and low fish intake had the lowest scores. The summary of this study is biased toward a description of the benefits aspect of this study rather than of the risk aspects. In fact, this study showed both benefit (increased VRM score) and risk (decreased VRM score) depending on the combination of fish intake and hair Hg concentration. This study points out the importance of the balance of omega-3 (as represented here by fish) consumption and MeHg exposure.

par. 3 and ff.

It is important to note that the Danish National Birth Cohort Study had no measures of MeHg exposure, and cannot, therefore, be used to evaluate the relation between risk and benefit.

p. 158 par. 4

This summary provides a very limited basis for comparing the specific intake of omega-3s and no basis for comparing the MeHg exposures across these studies. That is because fish intake *per se* is neither a common metric, nor a quantitative descriptor.

p. 159 par. 1

"…To express the results in terms of maternal DHA intake, one can assume that a women consuming more than 240 g of fish a week consumes white fish, oily fish and shell fish <u>in equal amounts</u>..." Why can one make this assumption? What is the basis for this? This assumption appears to be entirely arbitrary.

par. 2

"This is considerably lower in magnitude than estimates derived for one g/week maternal DHA intake in observational studies, approximately 1.38 verbal IQ points in Hibbeln et al. and 3.1 WEAMVA points in Oken et al. What is the basis for directly comparing IQ points and WRAMVA points? This has not been established in the draft Omega-3 Report.

par. 4

"However, the 4.1point higher…IQ scores of the children of fish oil supplemented mothers supports the plausibility of measureable neurodevelopmental benefit of maternal seafood consumption and gives one example of magnitude of dose-response. This conclusion confuses the issue. The 4.1 point average increased IQ score was derived from a study of fish oil supplementation, but is applied in the review to fish consumption *per se*. Because fish intake includes MeHg, fish oil intake and fish intake are not equivalent even if they represent equal fish oil intake.

p. 160 par. 2 and ff.

The inconsistency across studies of the outcome of fish oil supplementation in randomized clinical trials points out the danger in assuming the consistency of benefit across endpoints and developmental periods. This is particularly important in comparing benefit to risk since neither is consistent across endpoints.

p. 161 par 5 and ff.

"In both cohorts, the positive association of neurodevelopment with mothers' fish intake was stronger with additional adjustment for maternal mercury exposure." It is important to point out that since neither Daniels et al., nor Oken et al. measured omega-3 intake, the results of these studies cannot be generalized to other cohorts or consumers within these studies consuming fish diets that may differ in their omega-3/MeHg ratios from the average (but unknown) ratio in these studies.

p. 162 par. 2

Of all the studies included in the draft Omega-3 Report, the only studies that provide data that can potentially be generalized beyond the study cohorts themselves for deriving risk-benefit relationships are the two studies very briefly mentioned here (Budtz-Jorgensen et al. (2007), and Myers et al. (2007) (Actually, Strain et al. (2008), is the more current and more appropriate reference for this analysis of the Seychelles data). After considering all of the studies in the draft Omega-3 Report, we are left, in the end, with the most relevant studies being given barely a passing mention.

p. 176

The citation of Hibbeln (2008) does not appear in either the References list of the draft Omega-3 Report, or in Pub Med.

Addendum 2: Recommendations for Changes to Section II pp 13- 15

Note:

- Black text is material retained from the FDA draft Risk / Benefit Report document
- ~~**Red strikethrough text**~~ is material from the original document being recommended for deletion
- **Red underlined text** is additions being recommended to the document
- ***Red italicized text*** is comments related to revisions

SECTION II:

EXPOSURE TO METHYLMERCURY IN THE UNITED STATES

[…]

(e) Are Concentrations of Methylmercury Increasing in Commercial Fish?

~~**Most commercial fish species sold in the United States are harvested from the open ocean or from aquaculture sites. Aquacultured fish tend to be raised and harvested quickly without much opportunity to accumulate methylmercury. Moreover, aquacultured fish are not usually the large predatory types of fish that accumulate methylmercury over time by eating other fish containing methylmercury.** ***[Not related to question of increasing methylmercury]***~~

It has been estimated that human activity contributes over half of the total amount of mercury that is entering the atmosphere annually (EPA 1997). Increases in concentrations of methylmercury are ~~**more**~~ **especially** likely to occur in the vicinity of ~~**population**~~ **pollution** sources, e.g., in bodies of water such as rivers downstream from certain types of mining operations, and in ~~**relatively small,**~~ enclosed bodies of water such as lakes, **where pollutants can accumulate** (EPA 1997). **Sunderland (2007) shows that size normalized concentrations of mercury in various tuna species are enriched in the Mediterranean Sea and Atlantic Ocean compared to the Pacific Ocean, which has lower mercury levels.** ~~**Limited data suggest that methylmercury concentrations in commercial fish have not increased or decreased over time.**~~ **The currently available monitoring data are insufficient to determine if the methylmercury concentrations in commercial fish have been increasing or decreasing.**

While most commercial fish eaten in the U.S. are marine fish coming from different oceans and estuaries around the globe, a fraction are from domestic freshwater environments. Numerous examples exist in freshwater and coastal marine systems where reductions in atmospheric loads have resulted in lower mercury levels in biota, including fish. For example, it has been found that mercury levels in fish from Swedish lakes have decreased with atmospheric loads (Munthe et al., 2004). Also, a recent article by Monson (2009) performed a trend analysis of the mean mercury concentrations of a standardized length northern pike (55cm) and walleye (40cm) in a set of lakes across Minnesota. In this study, they found that there was a decreasing trend in fish tissue mercury concentrations before the mid-1990s, after which the trend reversed and an increase in concentration was documented. It is unclear what caused the observed trends, but the researcher suggest that the earlier

decrease may have been from reductions in local emissions, while later increases may be caused by increased global anthropogenic emissions. Other possible influences that were suggested by Monson include global climate change or changes in sulfate deposition, affecting mercury methylation rates.

Studies of **museum specimens of marine** fish, including tuna and swordfish**,** that were **preserved** up to 90 years ~~**old**~~ **ago** (Miller et al., 1972; Barber et al., 1972) report **mercury** levels ~~**consistent with**~~ **similar to** today's levels. In ~~**both**~~ **these two** studies the researchers discounted the possibility that these findings could have been affected by the preservatives used to store the fish as well as other conditions of storage, although the researchers in one of the studies admitted that the possibility could not be "rigorously excluded" (Miller et al., 1972). In another study that focused on conditions of preservation, however, the researchers concluded that, depending on circumstances, preservation techniques could substantially alter heavy metal concentrations in museum specimens of fish (Gibbs et al., 1974). For this reason, comparisons of contemporary fish to museum specimens should ~~**not**~~ be regarded as ~~**definitive**~~ **inconclusive**.

In a more recent timeframe, mercury concentrations in Yellowfin tuna caught off Hawaii in 1998 were found to ~~**be essentially identical to**~~ **not significantly differ from** those caught in the same area in 1971 – a span of 27 years (Kraepiel et al., 2003). ~~**The researchers engaged in "mercury biogeochemistry" modeling for the equatorial and subtropical Pacific in an effort to explain why these fish showed no increase in methylmercury in spite of increases in global mercury emissions over the past century. The most likely explanation, they concluded, is that mercury is converted into methylmercury (the form of mercury in fish) in the deep ocean, with transfer to the upper layer of ocean taking a minimum of 400 years. They noted that Yellowfin tuna and their prey swim in the upper layer. The researchers assumed that the total mercury concentration in the upper ocean layer had doubled between 1860 (the onset of the industrial revolution) and 1990. Nonetheless, that mercury would not convert to methylmercury or be absorbed by fish in the upper layer unless it first sank into the deep ocean and then circulated back over a long period of time.**~~ **The failure to detect a statistically significant increase in the methylmercury content in these samples, despite known global mercury emission increases, led the researchers to hypothesize that methylation of mercury might primarily occur in deep ocean sediments, requiring centuries of oceanic cycling to enter marine food webs. Recent empirical findings, however, (Ekstrom et al., 2006; Kirk et al., 2008; Monperrus et al., 2007a; Monperrus et al., 2007b; St. Louis et al., 2007; Sunderland et al., 2009, in press) demonstrate that mercury methylation *can and does* occur in the marine water column at ecologically significant rates. The failure to detect changes in the methylmercury content of particular fish samples may be more simply explained by Kraepiel et al.'s (2003) methodological constraints in methylmercury detection or anomalous characteristics of the Yellowfin tuna samples. In addition, the historical significance of anthropogenic mercury sources in eastern North America and Europe compared to a more recent rise in sources on the Asian continent may help to explain a relative enrichment of Hg in the Atlantic Ocean compared to the Pacific Ocean. However, recent modeling (Sunderland et al., 2009, in press) suggests that anthropogenic sources are now beginning to impact basin-wide concentrations in the North Pacific.**

~~**Mercury concentrations in freshwater commercial species are low. In our database the average mercury concentration for commercial freshwater species is 0.08 ppm on a per species basis, and the highest average for any species is 0.14 ppm (FDA 2006). (Recall that the average for all commercial species, weighted for consumption, is 0.086 ppm.)**~~ ***[Not related to question of increasing methylmercury]***

FDA's ~~**methylmercury**~~ database **of methylmercury levels in commercial fish** was reviewed for evidence of increases in concentrations over time. The database spans 30 years, starting around 1974. As described previously, for each species it includes the range of concentrations in the samples from highest to lowest and the mean concentration. For some species the database only includes recent sampling because interest in that species has been recent; for others the data span 20-25 years of sampling and for others the data span about 30 years. Overall, the database does not reveal a trend one way or the other, although the size of the database and the timeframes of collection are limited.

References Cited in Suggestions for Revisions

Cossa, D., Averty, B., Pirrone, N. (2009). The origin of methylmercury in open Mediterranean waters. *Limnology and Oceanography*, in press.

Ekstrom, E., Malcolm, E., Schaefer, J., Morel, F. (2006), Biotic and abiotic sources of methylmercury in the ocean, in *8th International Conference on Mercury as a Global Pollutant*, edited, Madison, WI, USA.

Kirk, J. L., V., Hintelmann, H., Lehnherr, I., Else, B., Poissant, L. (2008). Methylated mercury species in marine waters of the Canadian High and Sub Arctic. *Environmental Science and Technology*, in press.

Knightes, C., Sunderland, E., Barber, C., Johnson, J., Ambrose, R. (2009). Application of ecosystem-scale fate and bioaccumulation models to predict fish mercury response times to changes in atmospheric deposition. *Environmental Toxicology and Chemistry, 28(4),* 881–893.

Kraepiel, A.M.L., Keller, K., Chin, H.B., Malcolm, E.G., Morel, F.M.M. (2003). Sources and variations of mercury in tuna. *Environmental Science and Technology, 37(24),* 551-8.

Monperrus, M., Tessier, E., Amouroux, D., Leynaert, A., Huonnic, P., Donard, O. (2007a). Mercury methylation, demethylation and reduction rates in coastal and marine surface waters of the Mediterranean Sea. *Marine Chemistry, 107,* 49-63.

Monperrus, M., et al. (2007b). The biogeochemistry of mercury at the sediment-water interface in the Thau Lagoon: Evaluation of mercury methylation potential in both surface sediment and the water column. *Estuarine, Coastal and Shelf Science, 72,* 485-496.

Monson, B.A. (2009). Trend Reversal of Mercury Concentrations in Piscivorous Fish from Minnesota Lakes: 1982-2006. *Environmental Science & Technology, 43(6),* 1750-1755.

Munthe, J., Fjeld, E., Meili, M., Porvari, P., Rognerud, S., Verta, M. (2004). Mercury in Nordic freshwater fish: An assessment of spatial variability in relation to atmospheric deposition. RMZ-Mater. *Geoenvironment,* 1239-1241.

St. Louis, V., et al. (2007), Methylated mercury species in Canadian high Arctic marine surface waters and snowpacks, *Environmental Science and Technology, 41,* 6433-6441.

Sunderland, E., Krabbenhoft, D., Moreau, J.; Strode, S., Landing, W. (2009). Mercury sources, distribution and bioavailability in the North Pacific Ocean: Insights from data and models. *Global Biogeochemical Processes*, in press.

Sunderland, E., Mason, R. (2007). Human impacts on open ocean mercury concentrations. *Global Biogeochemical Cycles, 21*.

Sunderland, E. (2007). Mercury exposure from domestic and imported estuarine and marine fish in the U.S. seafood market. *Environmental Health Perspectives, 115(2),* 235–242.