Quantitative Assessment of the Net Effects from Eating Commercial Fish on Fetal Neurodevelopment (As Measured by IQ and also by Early Age Verbal Development in Children): Peer Review Report

May 2014

Introduction and Summary of Contents

During the summer of 2008, seven individuals expert in a range of scientific disciplines, identified below, completed an external peer review of a Food and Drug Administration (FDA) draft document subsequently published under the title "Report of Quantitative Risk and Benefit Assessment of Consumption of Commercial Fish, Focusing on Fetal Neurodevelopmental Effects (Measured by Verbal Development in Children and on Coronary Heart Disease and Stroke in the General Population." The peer reviewers were asked to provide individual, written comments in response to a specific set of questions and to provide any other comments on the draft they deemed appropriate. FDA made revisions to the draft assessment in response to peer review comments and then also in response to comments provided by other Federal agencies before publishing it for public comment in 2009. During that public comment period it was sent back to the peer reviewers for a second round of review, primarily to obtain their views on how FDA had addressed their original comments and to obtain any further recommendations they might have. This peer review report contains their comments verbatim followed by responses from FDA on those comments that requested changes or clarifications.

FDA is issuing this peer review report now to accompany the public release of the current version of that assessment, now entitled "A Quantitative Assessment of the Net Effects of Commercial Fish Consumption During Pregnancy on Fetal Neurodevelopment (As Measured by IQ and also by Early Age Verbal Development in Children)." The assessment describes quantitative estimates of the net effects on fetal neurodevelopment in children from maternal consumption of commercial fish, taking into account both adverse effects from methylmercury and beneficial effects from fish, presumably from one or more nutrients in the fish. It is narrower in scope than the version of the assessment last reviewed by the peer reviewers in that it no longer includes an assessment of the effect of fish consumption on risk of fatal coronary heart disease and stroke. That part of the assessment remains in draft as published in 2009. While the verbatim peer review comments contained in this report include those relating to the assessment of risk of fatal coronary heart disease and stroke, FDA is only responding at this time to those comments germane to the assessment of fetal neurodevelopment. The comments on coronary heart disease and stroke are being held in the event that that portion of the assessment is completed.

We greatly appreciate the peer reviewers' comments and suggestions. Since this peer review is a follow-on to the first peer review provided by the same individuals, the charge questions to the first peer review and to this follow-on review are both set forth below. The reviewers' full responses to the follow-on charge questions and their other comments are provided without attribution to the specific reviewer. The last section of this document identifies issues raised by the peer reviewers and sets forth whether FDA agreed or disagreed with the comments and describes the actions the agency took in response.

The Peer Reviewers (as of the time of the peer review)

Elaine M. Faustman, Ph.D.
School of Public Health and Community Medicine
University of Washington
Seattle, WA 89105

Dr. Faustman graduated *cum laude* with a dual major in Chemistry and Zoology from Hope College, Holland, MI and received a Doctorate in Pharmacology and Toxicology from Michigan State University. In addition to serving as a Professor in the Department of Environmental and Occupational Health Sciences at the University of Washington, Dr. Faustman is also the Director of the Institute for Risk Analysis and Risk Communication in the School of Public Health and Community Medicine, Director of the Center for Child Environmental Health Risks Research, Director of the Pacific Northwest Center for Human Health and Ocean Studies, and Director of the Reproductive and Developmental Toxicology Core, NIEHS Center for Ecogenetics and Environmental Health, and chairs the University of Washington Chemical Hazards Advisory Committee. She also serves as an Adjunct Professor in the Evans School of Public Affairs at the University of Washington and as an Adjust Professor in the Department of Engineering and Public Policy at Carnegie-Mellon University in Pittsburgh, PA. Dr. Faustman is a Fellow of the American Association for the Advancement of Science, a Diplomate of the American Board of Toxicology and serves as a panel member on the National Advisory Panel for the National Oceanic and Atmospheric Administration's (NOAA) Oceans and Human Health Initiative. Dr. Faustman's primary areas of study are the mechanistic investigation of reproductive and developmental toxicants; molecular mechanisms of action of metals and pesticides; quantitative risk assessment; development of biologically based dose response models for non-cancer risk assessment; in vitro toxicology; molecular epidemiology; toxicogenomics; and public policy. She has served as a presenter and/or chair at several international seminars and symposia, including the International Society of Exposure Analysis (ISEA), International Society for Environmental Epidemiology, the 5th Congress of Toxicology in Developing Countries, and the South Africa Toxicology Society. Dr. Faustman serves as a reviewer for a variety of scientific journals, including: *Applied Occupational and Environmental Hygiene Journal; Teratogenesis, Carcinogensis and Mutagenesis; the American Journal of Epidemiology;* and *Aquatic Toxicology*. She has published extensively on toxicity and teratogenicity of metals.

Herman J. Gibb, Ph.D.
Sciences International
Alexandria, VA 22314

Dr. Gibb received a Masters in Public Health (Environmental Health) from the University of Pittsburg and a Doctorate in Epidemiology from Johns Hopkins University. He is President of Sciences International. Dr. Gibb is an Adjunct Professor at the George Washington University School of Public Health and belongs to the International Society of Environmental Epidemiology. Dr. Gibb chairs the World Health Organization's Foodborne Epidemiology Reference Group (FERG) task force on foodborne chemicals and is a member of the FERG source attribution task force to evaluate methods to determine the disease risks from food versus those from water. Before joining Sciences, Dr. Gibb held positions as the Associate Director for

Health and Assistant Center Director at the National Center for Environmental Assessment of the U.S. Environmental Protection Agency (EPA). As the Associate Director for Health, Dr. Gibb was responsible for the Integrated Risk Information System, EPA's on-line system of health risk assessments. He was the Project Officer for EPA's cooperative agreements with the World Health Organization. He directed EPA's assessment of inhalation exposures and potential health risks to the general population that resulted from the collapse of the World Trade Center towers. He was the recipient of the EPA's Scientific and Technological Achievement Award for his study of lung cancer mortality and clinical irritation among chromate production workers and the recipient of the EPA's Gold Medal for Exceptional Service for his work on the drinking water standard for arsenic. Dr. Gibb was a member of the White House Interagency Committees on Mercury and Risk Assessment. He was the lead author of EPA's Mercury Research Strategy.

Dariush Mozaffarian, MD, Dr.PH, FACC, FAHA
Harvard School of Public Health
Boston, MA 02115

Dr. Mozaffarian received a medical degree from Columbia University, a Masters in Public Health (Epidemiology) from the University of Washington, and a Doctorate in Public Health (Epidemiology) from Harvard University. He is an Assistant Professor, Department of Medicine, Harvard Medical School and Assistant Professor, Department of Epidemiology, Harvard School of Public Health. He is founder and co-director of the Program in Cardiovascular Epidemiology at the Harvard School of Public Health. Dr. Mozaffarian's teaching has included Cardiovascular Epidemiology at the Harvard School of Public Health and Tufts School of Medicine. Dr. Mozaffarian's primary area of study has been the effects of lifestyle factors on multiple endpoints, including coronary disease, sudden death, stroke, heart failure, and atrial fibrillation. Examples of specific research projects include investigation of relationships of different fish meals with arrhythmic and non-arrhythmic coronary events; relationships of dietary fiber from fruit, vegetable, and cereal sources with stroke and coronary event; relationship of trans fatty acid intake with systemic inflammation; relationship of fish intake with heart failure and atrial fibrillation; effects of fish oil on heart rate in randomized trials, and mercury, selenium, and risk of cardiovascular disease in women and men. The aim of the latter research is to investigate prospectively the relationships of mercury and selenium levels and fish and omega-3 fatty acid intake with incidence of coronary heart disease and stroke. Dr. Mozaffarian has participated on a FAO/WHO Expert Consultation on Fats and Fatty Acids in Human Nutrition and on a United States Department of Agriculture Seafood Education Project Advisory Group. He has published extensively on the effect of fish and omega-3 fatty acid consumption on risk of coronary heart disease and stroke.

Gregory M. Paoli, M.A.Sc.
Risk Sciences International, Inc.
Ottawa, ON, Canada, K1N 6Z4

Mr. Paoli has a Master of Applied Science in Systems Design Engineering from the University of Waterloo. He is President of Decisionanalysis Risk Consultants, Inc., specializing in risk assessment and risk management in the field of public heath and safety. Within Canada, Mr. Paoli has served on Expert Committees of the National Roundtable on the Environment and the

Economy and is a member of Health Canada's Expert Advisory Committee on Antimicrobial Resistance Risk Assessment. He has provided guest lectures at the Queen's University's Public Sector Executive Programme and School of Public Policy, the University of Calgary's Faculty of Management and the University of Ottawa's Institute of Population Health. In the United States, Mr. Paoli has served on an Institute of Medicine Committee tasked to Review the United States Department of Agriculture's E. coli 0157:H7 Farm-to-Table Process Risk Assessment. He was appointed to a NRC Committee entitled "Improving Risk Analysis Approaches Used by the Environmental Protection Agency." Mr. Paoli served for several years on an Expert Panel to develop a risk ranking framework for the FDA and was on the peer review panel for the Harvard BSE risk assessment. He has served on several international expert panels including Expert Consultations as part of the Joint Food and Agriculture Organization and World Health Organization (FAO/WHO) Activities on Microbial Risk Assessment. Mr. Paoli has provided training in risk assessment approaches in North America, Japan, and South America. He also provides lectures as part of the Harvard School of Public Health continuing education course in Probabilistic Risk Assessment.

Barbara Petersen, Ph.D., M.P.H.
Exponent
Washington, DC 20036

Dr. Peterson received a Masters in Public Health in Nutrition from the University of California at Los Angeles and a Doctorate in Biochemistry from George Washington University. She is employed by Exponent, an engineering and scientific consulting firm, where she serves as Principle Scientist in Exponent's Health Sciences Center for Chemical Regulation and Food Safety. Dr. Peterson's areas of expertise include exposure assessment methodology, functional food safety and efficacy evaluations, food consumption profile modeling, and applications of Mote Carlo techniques to risk assessments for chemicals, including contaminants, pesticides and nutrients. Dr. Peterson has directed the design and conduct of numerous statistically based market basket studies, including acute and chronic assessments for pesticides, compliance assessments under proposition 65, and market research. Dr. Peterson served on the EPA Science Advisory Board's Integrated Exposure Committee and as an Expert Advisor to the FAO/WHO for several sessions of its Joint Expert Committee on Food Additives and Contaminants and for numerous consultations on risk assessment. She also served as Principal Investigator for the National Cancer Institute's International FOODBASE project, a major effort to collect and computerize descriptive and summary information on food consumption surveys conducted in more than 40 countries. Dr. Peterson has provided statistical support to FDA's Center for Food Safety and Applied Nutrition, including developing criteria for evaluating nutrition databases, and to EPA's Office of Research and Development. She has been a faculty member in risk assessment training programs for government scientists in the European Union, Thailand, the United States, and China. She has published extensively on methods for estimating dietary exposure.

Kimberly M. Thompson, Sc.D.
Harvard School of Public Health
Boston, MA 02115

Dr. Thompson has a Master of Science degree in Chemical Engineering Practice from the Massachusetts Institute of Technology and a Doctor of Science degree in Environmental Health from Harvard University. She is an Associate Professor of Risk Analysis and Decision Science at the Harvard School of Public Health Department of Health Policy and Management and the Department of Society, Human Development, and Health. She is the creator and Director of the Kids Risk Project. Dr. Thompson is also Associate Professor of Risk Analysis and Decision Science (Pediatrics), Children's Hospital, Harvard Medical School, where she is co-founder of the Center on Media and Child Health. Prior academic appointments include Visiting Associate Professor, MIT Sloan School of Management. Dr. Thompson's research interests and teaching focus on issues related to developing and applying quantitative methods for risk assessment and risk management, and consideration of the public policy implications associated with including uncertainty and variability in risk characterization. She is particularly interested in issues related to variability in risk for sensitive sub-populations, particularly children, and the potential risk tradeoffs associated with policies designed to protect them. The work includes research on a range of children's risks including injury, environmental, medical, and product-related risks, as well as perception of children's risks and the portrayal of risky behaviors in popular entertainment media. Her publications includes the book entitled "Risk in Perspective: Insight and Humor in the Age of Risk Management," a guide to help consumers take charge of health information.

Renee C. Wachtel, MD
Department of Developmental and Behavioral Pediatrics
Children's Hospital and Research Center at Oakland
Oakland, CA 94609

Dr. Wachtel received a medical degree from the State University of New York, Downstate Medical Center. She is Director of the Division of Developmental and Behavioral Pediatrics, Children's Hospital and Research Center at Oakland, Oakland, California. Previously, she was Professor of Pediatrics, University of Maryland School of Medicine, and the Director of the Division of Behavioral and Developmental Pediatrics, University of Maryland School of Medicine. She was also an Associate Professor of Pediatrics at The Johns Hopkins School of Medicine. Dr. Peterson serves as Co-Chairperson, Committee on Developmental and Behavioral Pediatrics, Northern California Chapter, American Academy of Pediatrics; Co-Chairperson, Committee on School Health, Northern California Chapter, American Academy of Pediatrics; and Chair, Autism Task Force, Children's First Medical Group. She has conducted research and published extensively on a range of neurodevelopmental issues.

The Charge to the Peer Reviewers (ORIGINAL)

Each expert peer reviewer was provided with a written "charge" concerning the document, as follows:

"When the FDA generates a scientific assessment, it is presenting its scientific evaluation about the accumulated evidence. The peer review should provide input on the reasonableness of judgments made from the scientific evidence. The result should be an independent determination by each peer reviewer as to the appropriateness of (a) the assumptions made and hypotheses postulated, (b) the methodology utilized, (c) the quality and relevance of the data and information, (d) the accuracy of the analytic results, and (e) whether the conclusions reached are supported.

Charge Questions

1. Is the document logical and clear?
2. Were scientific assumptions explained and are they appropriate?
3. Has the appropriate literature been cited? Are there publicly available, peer-reviewed papers that should be included?
4. Do the conclusions follow from both the analysis of the studies that are reviewed from the peer-reviewed literature and from the results of the quantitative risk assessment?
5. Specifically in regards to the quantitative risk assessment:
    1. (5a) Were sufficient information and explanations given that describe how the data were identified and what criteria were used to determine the suitability of the data? Were these criteria adequate? Was the methodology appropriate? If there are other data that should be included in the quantitative analysis, how should it be used?
    2. (5b) Is the model clearly described and is it supported by existing data? Are uncertainties in the model identified and characterized? In particular, does the uncertainty analysis encompass the range of plausible assumptions?
    3. (5c) Are scientific assumptions explained and are they appropriate?
6. Are there additional endpoints of risks associated with methylmercury or of benefits of fish consumption that were not modeled and should have been? If so, what are they and what data should be used?
7. Are the intervention scenarios appropriate?"

The Charge to the Peer Reviewers (FOLLOW-ON REVIEW)

"The purpose of this charge is to request any follow-on written comments by the external peer reviewers on the new draft reports:

- "Report of Quantitative Risk and Benefit Assessment of Consumption of Commercial Fish, Focusing on Fetal Neurodevelopmental Effects (Measured by Verbal Development in Children) and on Coronary Heart Disease and Stroke in the General Population"[1] and

- "Summary of Published Research on the Beneficial Effects of Fish Consumption and Omega-3 Fatty Acids for Certain Neurodevelopmental and Cardiovascular Endpoints"

[NOTE: The summary of published research remains in draft form as published in 2009. FDA has not made a decision on whether to update it. The summary served an important function in 2009 but an update to it is not essential at this time.]

The initial draft report has now been divided into two separate documents *[but see NOTE, above]* and substantial revisions have been made to the text previously reviewed by the external peer reviewers, both in response to their comments and pursuant to comments received in inter-agency review. Any additional, follow-on comments should be confined to addressing the following questions:

1) Do you think the creation of two separate documents enhances or distracts from communication of the information contained in the report?

2) Do you have any comments on the revisions made to the draft assessment?

3) Do you have any additional comments on this draft assessment?"

---

[1] As stated previously, the name of the assessment has been changed. Its name as published in draft in 2009 was as stated in this "bullet."

The Peer Reviewer Comments to the Follow-On Charge (provided in random order and without attribution)

Reviewer #1

I. General Impressions

The readability of the document has been significantly improved. The new executive summary is very good and accurately captures the conclusions of the assessments.

II. Response to Charge Questions

1) Do you think the creation of two separate documents enhances or distracts from communication of the information contained in the report?

Yes, this is a major improvement. Both documents make valuable contributions. Readers will find the two-document format much easier to use. I was able to much more readily discern what data had been used, and how it had been used, in the risk assessment in this new format.

2) Do you have any comments on the revisions made to the draft assessment?

It appears that FDA has seriously considered all of the reviewers' comments and taken on board those that could be incorporated. The table explaining their decisions was particularly helpful to me. I think their responses were appropriate given the available data and the purposes of the assessment.

3) Do you have any additional comments on this draft assessment?

Yes. I believe FDA misunderstood one of my comments regarding QA/QC. My comment refers to the technical checks of each calculation rather than the peer review of the methodology. That is, I was commenting on the need for a detailed (and independent) check of the models and data in addition to the peer review. In other words, I agree with the selection of the data set, models and the apparent application of the models to the data…but I have largely assumed that the software has actually done the calculation as described. My concern is that these are extraordinarily complicated, data intense models and there needs to be independent confirmation that the actual calculations have occurred as described (by independent I would suggest that another FDA staff member randomly check different components of the model).

III. Specific Observations

The document needs a glossary of terms/definitions and as detailed an index as possible.

Reviewer #2

I. General Impressions

Please see responses to charge questions, below.

II. Response to Charge Questions

1) Do you think the creation of two separate documents enhances or distracts from communication of the information contained in the report?

In my opinion, the separation into two documents greatly enhances the readability of the document. Separating the review of published research from the assessment of risk and benefit enables the reader to better understand the quantitative risk assessment, with easy access to the research upon which it is based in the separate volume. Both volumes would be much more useable if each had a table of contents. The summary tables in the research review volume also add to the ease of information access and easily enable the reader to compare research studies on relevant variables.

2) Do you have any comments on the revisions made to the draft assessment?

Yes. Section III A "Scientific basis for risk and benefit assessment" p 16-27 is much improved. It is clearly written, and conveys good information.

There are extensive revisions to Section V, some of which clarify the assessment and others that are still not sufficiently clear. For example, the discussion on page 85 about serving sizes would benefit from including ounces or servings in Table V-1, in addition to grams which are not generally used by American consumers.

In the discussion of modeling on pages 87-103, the explanation of Z scores and IQ size equivalents is fairly clear. While IQ scores have been repeatedly shown to follow a normal distribution, I question whether there are data to show that age of walking and age of talking follow a normal distribution, and I know that the means and medians of these milestones vary in normal populations of different ethnicities. In fact, the average age of walking unaided in the United States is now 11.5 months, but the Iraqi study (Marsh 1987) had only 25% who walked at 12 months or less, and does not reference population norms. The statement on P 68 that "milestone standard deviations to not vary greatly among populations" is not referenced, and I do not know its accuracy for age of talking. I also think that converting ages of milestone attainment into "IQ se" is very questionable from a clinical and methodological perspective.

I have serious concerns about taking the delays in talking and walking, which was recorded in months, and using statistics to show differences in hours, which is way beyond the accuracy of the data on which it is based (see below). I strongly suggest that they be eliminated from Tables V4 and V5.

3) Do you have any additional comments on this draft assessment?

Yes, I have a number of comments:

1. There are a number of typos or inconsistent attributions in the Executive Summary (e.g. "National Academies of Science" versus "National Academy of Sciences;" "IOM 2006 at 1 or 6" (???) both on Page 2; reference "Willats et al 1989" page 19).

2. I appreciate the novel approach of the FDA risk/benefit assessment in relation to neurodevelopment, and the expertise of the FDA's statistical team. However, as an expert in pediatric neurodevelopment, I continue to have serious concerns about the choice of two specific developmental milestones (age of talking and age of walking) as primary indicators of the effects of MeHg on fetal neurodevelopment, especially when there is much better data about developmental outcomes from other studies, for example the studies by Oken et al 2008 and Hibblen et al 2007. The age of walking and first talking is an extremely crude measure of child development, even if the data were extremely accurate, and do not evaluate a whole range of more complicated developmental skills that could be impacted by maternal seafood consumption during pregnancy. I think the discussion of IQ outcomes, e.g. Table V6, is much stronger information.

3. I also continue to have concerns about the methodology employed to assess risk/benefit (see above). The Iraqi study (Marsh 1987), upon which much of this analysis is based, had great imprecision in data collection about development. Mothers "did not know exact birth dates." The mean age of the child when first seen was 30 months (with some much older), and mothers were asked to recall when their child first walked unaided or spoke "2-3 meaningful words." At very best, these estimates "may have been inexact," in the words of the authors of the studies. Therefore, any use of this data must be very cautious, and not imply any greater precision that the original data can support. Therefore, in my opinion, the unit of months is the highest precision this data would allow, not days or hours.

4. I feel that while the report does not focus on ethylmercury, many consumers reading the report, particularly the executive summary, do not understand that this report does not address neurodevelopmental effects of ethylmercury, and would benefit from a statement such as "This study does not address the potential neurodevelopmental effects of ethylmercury, and the relationship of mercury exposure to autism." I also take issue with the footnote on p 10, which states that exposure to thimerosal is "extremely small, occurring once per year at most." Until recently, American infants were given multiple doses of thimerosal containing vaccines during the first year of life, and may have been given these vaccines at a time when the brain might have been particularly vulnerable to ethylmercury effects.

5. The discussion about the comparability of the age of talking with the MacArthur Communicative Development Inventory (MCDI) and the Denver Developmental Screening Test (DDST), p71, misses two important points. Language development has both a receptive component and an expressive component. Age of first talking only addresses the expressive component. Both the MCDI and the DDST address both receptive and expressive language. Both are tests that are standardized, and used for screening. The DDST does not give age equivalents,

only pass or concern. So it is not clear to me how the data were used, and if it was appropriately combined in the model.

6. The definition of "age of first talking" is not clearly defined in the report, and this will make most readers confused about the resulting risk/benefit assessment. As I stated in my previous comments, speaking 2-3 words can be normally at 11 months if you include "mama" and "dada" or other equivalent names, and 13-15 months if you exclude names.

7. I am surprised that there is no information in Table AA-2 about shrimp, which is 22% of market share, or canned salmon

8. Table V-10 should consistently use "IQ se," not "IQ point" since they are not equivalent.

Finally, the report addresses important health concerns for the American public in general, and for health professionals who advise women in their childbearing years. A one paragraph abstract, which could be disseminated as part of health promotion by health care providers, stressing the positive benefits of eating low MeHg containing seafood, would facilitate distribution of this important message

III. Specific Observations

Please see responses to charge questions, above.

Reviewer #3

I. General Impressions

Overall, the changes to the report improve its clarity and presentation. I still believe, however, that the entire report would benefit from substantial editing and that this will significantly improve the clarity of the presentation. The report reads much better than the earlier draft, but I found numerous typos and it was clear that the report still has not been rigorously or professionally edited. As a scientific report, it is complete and I can generally figure out what the authors have done, but it is much more tedious to read than it needs to be, and I found it frustrating as a reviewer to be noting many minor punctuation and grammatical errors (e.g., some sentences are missing periods while
others have two; hyphens are missing in many places). One aspect that I believe a good editor would address is the lack of consistency in the terms used to refer to the risk and benefit assessment itself, which gets referred to as the: "risk and benefit assessment," "risk and benefit assessment report," "benefit and risk assessment report," "risk-benefit assessment," "risk/benefit approach," "risk and benefit analysis," etc. Now that I see it in the draft, I think I would prefer consistent use of the term "benefit-risk assessment" or "risk-benefit assessment" throughout, but my point is that the authors should choose one term and use it consistently throughout. I appreciate that the authors added the word "benefit" to the title and throughout as I suggested in my last set of comments.

I think that the reorganization has improved the report significantly, but I still think that a better roadmap at the beginning of the Report and a section of conclusions or insights (e.g., a short new Section VI starting on page 115 before Appendix A) at the end would be helpful. I fully support the separation of the summary of the published literature into a separate document, which makes it much easier to get a better sense of the weight-of-the-evidence and helped to consolidate the risk benefit assessment information into a report of appropriate size. I also found that the addition of the figures in Section IV were very helpful, particularly Figures IV-2 and IV-3 and their associated tables. I have again made extensive suggestions using track changes instead of trying to itemize them here.

II. Response to Charge Questions
.

1) Do you think the creation of two separate documents enhances or distracts from communication of the information contained in the report?

My view is that this separation significantly enhances the communication of the information related to the risk and benefit assessment.

2) Do you have any comments on the revisions made to the draft assessment?

I commend the authors on their comprehensive responses to the comments. I believe that they made appropriate revisions in response to my comments, except that the current version still needs editing.

3) Do you have any additional comments on this draft assessment?

I think that this is a very important report, and one that I believe will improve understanding of the risks and benefits of consuming fish.

III. Specific Observations

See attached with track changes. I have focused my comments on the benefit-risk assessment and I am not providing any specific comments on the companion summary document (please note however, that this document could also benefit significantly from professional editing).

A track changes version of the report has been provided to CFSAN authors for consideration.

Reviewer #4

I. General Impressions

The USFDA scientists deserve a hearty round of applause for the excellent job they have done on these reports. I was encouraged to see the extensive response to reviewer comments and major changes taken in this document. As a reviewer, I felt that the FDA scientists responded in a genuine and scientifically defensible manner to the issues raised. I have provided some

comments on the specific responses, as well as responded to the three general charge questions below.

In the numerous cases where this reviewer did still have issues of interpretation and issues with extrapability of report comments, this reviewer feels that the use of sensitivity analysis and decision analysis approaches by FDA scientists led to an informed discussion. In some cases, this reviewer continued to have a difference of scientific opinion that was both natural and expected in a lengthy and very detailed report such as was prepared.

This reviewer has several comments on how changes could have been documented. It is extraordinarily easier for reviewers to follow changes if the use of "track changes" is done in order to see how these long and detailed documents are modified. In reviewing these documents, this reviewer checked general concepts but did not review line by line the numbers nor could this reviewer determine if significant new model runs were conducted in response to the collective reviewer comments. In fact, even when the Authors state that they have conducted more analysis (as was discussed with response 1 on Page 53 of responses to reviewers), this reviewer could not easily find the discussion of these new analyses. Despite these limitations, this reviewer feels that significant progress was made in addressing the reviewer comments and in developing a readable and highly useful document.

II. Response to Charge Questions

1) Do you think the creation of two separate documents enhances or distracts from communication of the information contained in the report?

This reviewer feels that the separation of the two documents significantly enhances the readability and understanding of the reports. Several additional changes should occur to support this separation. A Table of Contents with the outline of the report is essential for both sections.

In several sections it was difficult to find the bottom line or summary of the preceding section. For example, in the "Report of Quantitative Risk and Benefit Assessment…," Appendix A starts on page 115, but where are the final summary comments for the text? The overall summaries are at the start of the document but some additions to the "end" of the report should also occur. Otherwise the reader is left without focus. Several formatting issues like this example could provide an additional level of enhancement.

2) Do you have any comments on the revisions made to the draft assessment?

Reviewer comments on FDA Response to the Peer review comments:

Draft Risk and Benefit Assessment Report Peer Review Report

1.   Page 50, question 4. This reviewer does not agree with response.

2.   Page 50, question 7. FDA scientist could think of an IEUBK like response for relative source contributions for commercial and non-commercial sources of fish.

3.   Page 51, question 8. In order to answer this question, shellfish eaten should be considered not just LOD values. Please be explicit, as <LOD doesn't mean there are no public health significant effects.

4.   Page 51, question 9. Yes, that is exactly the point made by the reviewer comments.

5.   Page 51, question 13. This reviewer continues to disagree with this approach.

6.   Page 51, question 14. This reviewer does not agree with FDA's response for this issue.

7.   Page 51, question 3. Good response.

8.   Page 53, question 6. Thank you for giving additional endpoint variability.

9.   Page 53, question 1. Thank you for doing this extra analysis.

10.   Page 54, question 3. Thank you for your additional considerations to this question, however, this reviewer's initial concerns still remain.

11.   Page 54, Use of data at later ages. Thank you for the additional modeling and conducting a sensitivity analysis.

12.   Page 55, question 1 and 2. This reviewer feels additional efforts could have been done in this area.

13.   Page 55, Genetic susceptibility to methylmercury. This reviewer does not agree with this response.

14.   Page 56, Postnatal exposure, questions 1-3. This reviewer does not agree with deleting the section on postnatal exposure.

15.   Page 56, Effects of other contaminants, questions 1-3. This reviewer feels strongly that these points and limitations need to be discussed in this report.

16.   Page 56-57, question 1. This reviewer feels that some greater context was needed for results on this topic.

17.   Page 58, "Roadmap" of the Contents: Good additions were identified.

18.   Page 60-61, question 1. Thanks for adding these percentages.

19.   Page 61, Account of Cardiovascular Disease. Great additions.

20.   Page 62, question 2. This reviewer feels authors should still acknowledge this potential co-variant for their assessment of low nutrient fish.

21.   Page 64, question 1-2. Thanks for adding these details and results.

22.   Page 65, question 5. Thanks for this addition.

23.   Page 65-66, questions 1 and 2. Thank you for removing these references.

24.   Page 67, Male vs. female data in the studies used for modeling stroke. Important for authors to have added this limitation.

25.   Page 69, Sufficiency of documentation on the technical detail of the modeling. Great response.

26.   Page 69, Appendices. Thank you for removing Appendices E through I.

27.   Page 70-73. This reviewer was pleased by the development of this separate report and by the author's response to reviewer comments.

Summary of Published Research on the Beneficial Effects of Fish Consumption and Omega-3 Fatty Acids for Certain Neurodevelopmental and Cardiovascular Endpoints:

Please see my general comments that I have made in response to Charge Question 1.

1.   Page 62, last line, typo. Remove 0.

2.   Page 101. Reviewer noted some distinct differences in formatting and summary approaches throughout this document. For example, see San Giovanni et al 2000a pg. 98-109 versus Simmer 2001 pg. 120-123.

Report Of Quantitative Risk And Benefit Assessment Of Consumption Of Commercial Fish, Focusing On Fetal Neurodevelopment Effects (Measured By Verbal Development In Children) And On Coronary Heart Disease And Stroke In The General Population:

1.   Page 16-17. Clear statement of purpose and context for the scientific analysis for Risk and Benefits. It appears that the authors have separated the analysis from claims about policy implications that were in previous appendices. Good revision.

2.   Page 39. The FDA scientist provided a much clearer indication of their conceptual model. See Figure IV.1 Basic Modeling Structure. Well done.

3.   Page 56, Figure IV-3: Flow Diagram for the Dose-Response Modeling and Table IV-2: Dose-Response: Limitations in knowledge, Assumptions that Address those Limitations, and Implications for the Results. Authors provided useful context for understanding analysis, approaches, assumptions, uncertainty and dose-response model approaches. Well done.

4.   Page 131-133, typographical problem was identified with mislocated figures and figure legends.

3) Do you have any additional comments on this draft assessment?

Again, this reviewer found these reports to be extremely useful, interesting and in some cases provocative, in a good manner. When published, this reviewer will use and refer to these documents. The missing element in both of these reports is a good list of "additional research needs" that would be targeted to the critical missing data throughout the report. The authors frequently referred to missing information and approached the impact of this missing data through sensitivity analysis, but what a "missed opportunity" to highlight the essential missing data in a concise, summary manner. Obviously, researchers always long for this section, but for these reports it is really such an excellent examination of the total data base on these topics that this really does seem to be a critical missing element.

III. Specific Observations

Please see responses to charge questions, above.

Reviewer #5

I. General Impressions

It is a difficult document to follow and requires several reads to understand what is going on, particularly with regard to the neurodevelopmental analysis. Much of that is simply the difficulty in imparting a cogent argument (assessment) in a lengthy document. For example, the "bottom line" on neurodevelopmental effects is based on early age verbal skills, but considerable information is provided on IQ (Axelrad 2007) and "age of walking" leading the reader to believe that they will eventually be considered in the overall model. It takes several reads to understand that is not the case. How the net effect is calculated is not even revealed until the first paragraph on page 94. The main text of the document should be in proportion to that which is essential to the bottom line, and the rest should be left to appendices (i.e., the document would benefit considerably by some tactical streamlining).

II. Response to Charge Questions

1) Do you think the creation of two separate documents enhances or distracts from communication of the information contained in the report?

I think that it enhances the communication of the information contained in the report. The former document had a considerable amount of information which I would not necessarily consider extraneous, but certainly not critical to the development of the analysis. The title of the second document, "Summary of Published <u>Research</u> on the Beneficial Effects of Fish Consumption and Omega-3 Fatty Acids for Certain Neurodevelopmental and Cardiovascular Endpoints" is misleading, however, since the document is not limited to research findings but

also contains reviews, meta-analyses and risk assessments. Perhaps the document should have subsections (e.g., government reports, original research, etc.).

2) Do you have any comments on the revisions made to the draft assessment?

The Executive Summary is improved in terms of length and clarity. The delineation of the assumptions are critical to the risk assessment, but the long tables on pages 46-48 and 57-65 interrupt the flow of the argument. It is recognized that the tables are intended to be read in tandem with the flow diagrams on pages 45 and 56, but many of the assumptions are claimed to have "negligible impact." Perhaps the assumptions which are believed to have the greatest impact on the assessment could be discussed in the text and the tables included as appendices? Tables IIIB, IV-3, and IV-5 should also be considered for the appendices.

3) Do you have any additional comments on this draft assessment?

The message of this document is that for the vast majority of the American population the benefits of eating fish outweigh the health effects of the methylmercury contamination. The criticism of the document will be that which is identified on page 94: "The disadvantage of this presentation is that, by necessity, it is limited to people who eat a variety of fish that, over time, contain both an average amount of methylmercury for commercial fish (0.086 ppm) and an average amount of nutrients that contribute to a beneficial net effect for fetal neurodevelopment." The document should acknowledge that it is possible that for regular consumption of certain kinds of fish, methylmercury could present a problem.

III. Specific Observations

Page 90: On page 54, it states that the model estimates that the age of first walking ranges from 6.3 months to 17.8 months with a median estimate of 10.4 months. 10.4 months is a young median age for walking. Estimates reported in the literature are about 12 months.

Page 54: It is unclear why mercury, other than methylmercury, would be excluded from the analysis. If mercury, other than methylmercury, is normally present in the body and it is believed to have similar health effects to that of methylmercury, then it would seem reasonable to include it in the analysis as a baseline of exposure. One of the peer review comments on the earlier document was that other exposures (e.g., PCBs) should be included in the analysis. The response to the comment was that such information was beyond the scope of the document, presumably because such information was not available and it would be difficult to combine results from MeHg and PCBs. But if the information on inorganic mercury is available from NHANES and its effects are believed to be similar to MeHg, why not use it?

Reviewer #6

I. General Impressions

Again, the U.S. FDA should be lauded for undertaking this crucial and critically important analysis and revision. Public health recommendations must have as their principal intent the improvement of health. For a food such as fish, considering both risks and benefits is essential to understanding health effects and making public health recommendations about consumption. The revision is considerably improved, and many of the prior limitations have been addressed. However, close attention to several remaining important issues is necessary so that this document will neither cause confusion nor lead to incorrect conclusions.

II. Response to Charge Questions

1) Do you think the creation of two separate documents enhances or distracts from communication of the information contained in the report?

Nice improvement.

2) Do you have any comments on the revisions made to the draft assessment?

1. Given the scientific and public health importance of these topics, the methods and assumptions for modeling the health risks and benefits must be transparent and replicable, particularly when different methods provide different results (for either central estimates or uncertainty of these estimates).

In particular, there is no rationale for the unusual methods for the CHD and stroke "pooled analysis models." Traditional random effects meta-analysis should be used, accounting for dose-response and both within-study and between-study variance and covariance. For such an important topic, this should be done using well-established and replicable methods, such as generalized least squares for trend estimation of summarized dose-response data (e.g., GLST in Stata). Having no assumption of one central "truth" is simply random effects meta-analysis. Accounting for dose-response and the appropriate variance-covariance network can be accomplished with GLST.

The data should then be evaluated <u>both</u>:

(1) Using all the available data simultaneously (one-stage random effects meta-analysis) as the principal analysis, and
(2) After first deriving each within-study dose-estimate (two-stage random effects meta-analysis) as a sensitivity analysis.
Both can be accomplished with the GLST command. The first estimate makes the most efficient use of the data and should be the primary analysis; the second is conservative and can be done as a sensitivity analysis. Weighting should be done by the inverse-variance of each stratum-specific RR or by the inverse of the number of events (cases) in each stratum, <u>not</u> be done by the size of

the population or the person-years of follow-up, as power (precision) is largely determined by the number of cases, not the size of the overall population.

The current opaque techniques used for the "pooled analysis" models can then be dropped. Such techniques were not applied by the authors to analyses of MeHg and neurodevelopment harm or of fish intake and neurodevelopment benefit. None of the stated reasons for using these unconventional techniques are sufficient. For example, the techniques are not justified by "the possibility that these populations have significant differences in terms of confounding." If any study were felt to be internally biased (confounded), nothing about the variance assumptions will alter or correct that. Not having an assumption of one central "truth" is fine, but this requires only traditional random effects meta-analyses.

Also, in conformance with standard meta-analyses, all data should be extracted independently in duplicate. A review of just a subset of the extracted data reveals several errors. For instance, in Table AA-15, the numbers of events listed for cohort #11 are for total CHD in that report, not fatal CHD. The characterization of data from cohort 12 also makes no sense: five different sets of what appear to be quintile data are shown, each with identical RR's and (very low) numbers of events. This study evaluated only two different types of fish meals separately, with many more events than are shown; this data would be better represented by two (not five) sets of quintile data for these two different types of fish meals, with actual reported RR's and numbers of events, and one overall estimate for the study.

This reviewer applied GLST to the data in Table AA-15 to estimate the RR of CHD death per 100 g/d of fish intake, after correcting the errors noted above and weighting each stratum-specific RR by the inverse of the number of cases in each stratum. Considering cohort #12 as two different studies (note this is not strictly correct), the one-stage GLST resulted in a RR of 0.61 (95% CI=0.58-0.64); note the narrow CIs that are concordant with this enormous body of data. The two-stage GLST (sensitivity analysis) resulted in a RR of 0.74 (95% CI=0.53-1.03), but this considers cohort #12 as two different studies. After requesting unpublished data from cohort 12 for the actual stratum-specific RR's of CHD death for overall fish intake, the one-stage GLST resulted in a RR of 0.58 (95% CI=0.53-0.65), and the two-stage GLST (sensitivity analysis) resulted in a RR of 0.68 (95% CI=0.49-0.94). Preferably, these analyses should be repeated using the actual SE of each stratum-specific estimate (rather than inverse-number-of-cases).

2. Page 35, page 77, page 101, Table V-11, etc. : The estimated numbers of CHD and stroke deaths prevented due to fish consumption in younger men and women are only correct given a very strong (and probably incorrect) assumption that fish intake at these younger ages has no effect on CHD and stroke deaths later in life. It is likely, based on the evidence for effects of fish consumption on CVD risk factors and the long time course of atherosclerosis and cardiovascular disease, that a substantial number of CHD and stroke deaths will be prevented at later ages due to fish consumption at earlier ages. In other words, fish consumption may have some acute benefits (e.g., on arrhythmia), but also has many other established chronic benefits (e.g., on blood pressure, inflammation, triglycerides, vascular and ventricular modeling). Thus, fish consumption at younger ages is likely to prevent a substantial number of cardiovascular deaths later in life. Available data does not let us directly estimate these numbers. Therefore, all of the

estimates in this report for CHD and stroke deaths prevented at ages 16-45 should include a prominent footnote/caveat that, due to the chronic nature of cardiovascular disease, fish consumption before age 46 is also likely to prevent some additional number of CHD and stroke deaths later in life.

3. What if scenarios: That only 1% of fish eaters might stop consumption is an unrealistically small unanticipated effect of any Advisory warning about fish consumption, no matter how focused.

3) Do you have any additional comments on this draft assessment?

Executive Summary: A few key summary tables and figures should be added.

Page 29, Table IIIB-2, etc. A key new meta-analysis should be added and discussed – a new pooled analysis of observational studies of fish consumption and primary prevention of coronary heart disease death, including 16 prospective cohort studies totaling 4,473 cardiac deaths in 326,572 generally healthy individuals (Harris WS et al. Towards establishing dietary reference intakes for eicosapentaenoic and docosahexaenoic acids. J Nutr. 2009;139:804S-19S).

Summary of Published Research, page 60, second paragraph: These critically important issues relating to underestimation of quantitative dose-response relationships derived from questionnaire estimates vs. biomarkers should be duplicated in the Executive Summary.

III. Specific Observations

Main Report

Page 16: Should state "scope of this report is limited to potential consequences to the developing nervous system in utero." The nervous system continues to develop for at least 2 years after birth.

Page 34: If the very limited data on MeHg and intermediate CVD risk factors such as blood pressure are going to be reviewed (that was against the stated aims of the report to focus only on disease endpoints), then the authors should review and include in detail the far more extensive data on fish or fish oil consumption and lowering of CVD risk factors, including blood pressure, heart rate, triglycerides, etc. (e.g., Am J Clin Nutr. 1997;65:1645S-1654S; J Hypertens. 2002;20:1493-1499; Circulation 2005;112:1945-1952; Am J Cardiol. 2006;97:216-222; etc.)

Tables AA-14 and AA-15. Titles should be changed to "coronary heart disease mortality," as cardiovascular mortality would include stroke, which is separately assessed.

Some stated interpretations of the wide "pooled analysis" CI's require correction. The explanation that "some people in the US could be experiencing circumstances similar to those in Eastern Finland" is not relevant to the CI's of the overall estimate. The upper bound of the CI here does not indicate that a small subset of the US population may have higher risk, but that the true net effect of overall fish consumption, based on all the world evidence, could be harmful for

CHD. This is clearly outside the range of believability of the worldwide data. Rather than using the unusual methods and assumptions of the CHD and stroke "pooled analysis model," not applied to any other modeling of other outcomes for fish or MeHg, the authors should use traditional and established meta-analysis methods (see above).

Responses to Reviewers

Page 67: If future sensitivity analyses are to consider CVD benefits using only data from US populations, then the same should be done for potential neurodevelopment and CVD harms of MeHg (i.e., use only data from US populations). If the authors are truly concerned about the validity of pooling data from different study populations, each with its own risk factors and applying the results the entire US population, then these same strict assumptions must be extended to all analyses. This reviewer would not advise such US-centric analyses – true biologic interaction is the exception rather than the rule, and it is unlikely that effects of fish, omega-3 fatty acids, or MeHg are biologically different in other countries.

Reviewer #7

I. General Impressions

I remain impressed by the overall effort in the assessment, particularly with the level to which raw data has been collected, scrutinized and re-processed to make the best potential use of the evidence base. This is particularly critical when the effects of interest are both subtle and diverse in their nature. The re-organized report further emphasizes the extent to which methods have been employed to formally process and draw inferences from the data, as a very distinct matter from the more standard descriptive review of evidence.

Despite the improved overall report, one negative impression that remains pertains to the overall lack of discussion on the overall level of confidence that should be asserted in the quantitative results. While some aspects of the uncertainty analysis are relatively explicitly detailed (e.g., explanations of how alternative fits of probability distributions to raw data are made, weighed and combined), the overall level of uncertainty at the level of model results is not described in a proportionate manner. One might expect considerably more discussion of uncertainty at the level of results, as compared to the uncertainty in any single element of the model. As an example, Appendix B indicates that modeled hair levels are between 30-40% lower than the NHANES data (p. 169). Two possible explanations are provided, but neither is argued to actually explain the discrepancy (it "may be explained, at least in part"). While there is some plausibility to the explanations, there does not seem to be any reflection or discussion on the possibility that the rather complex exposure modeling process is systematically underestimating exposures (which would be a straightforward conclusion unless one of the possible explanations is asserted as being very likely to account for the difference). There also doesn't seem to be any reflection of this potential underestimation in the overall process of providing confidence intervals in final results. This is just one example of an unfortunate imbalance between detailed characterization of uncertainty at the micro-level (e.g., we know exactly how an individual's exposure is constructed from a probability tree of possible distributions fitted to Hg measurements for individual species) and characterization of uncertainty at the macro-level (i.e., given all of the

quantified and unquantified uncertainties, what can be concluded about specific quantitative estimates, as compared to what is reflected in the provided confidence intervals). Perhaps the most glaring example of this imbalance is the provision of many tables of confidence intervals for both the pooled model and the meta-analysis model. On one hand, the analysts should be complemented for having provided both of these sets of estimates as a transparent display of the impact of a key uncertainty. The authors should in no way be faulted for not being able to generate a single clear statement of risk (or a single theory of uncertainty) from the nightmarish inferential challenge that is clearly demonstrated in Figure AA-15 in the data on CHD mortality. However, they seem to address the challenge too indirectly, by providing tables of results with both sets of estimates, rather than more directly discussing the implications and providing a narrative summary of the implications of this key uncertainty on the estimates (a few scattered lines on pages 102-105 suggesting that "it is more likely than not that increased fish consumption leads to a decrease in CHD death.").

A key aspect of a more complete narrative on the uncertainty analysis would be to provide the reader with some guidance in interpreting confidence intervals that span a continuum from one extreme, appearing to have a significant protective effect, to the opposite extreme of appearing to generate additional mortality. For CHD mortality in the pooled estimate, the potential for reversal of causality appears to occur at approximately the 85%ile (p. 106) of the uncertainty distribution. Some readers may take this to conclude that the net result of the analysis is that the direction of the causal relationship between fish consumption and CHD mortality is presently best described as unknown, even at the population level. If the fundamental question of causality is still in play, then this should be made much more apparent and not embedded in a confidence interval. If the causality is not really in question (e.g., whether it be from RCTs or another source of evidence not employed in generating these confidence intervals), this should be made clear.

The use of language may also be seen as somewhat biased in the author's attempts to diminish the impact of the pooled analysis model (e.g., in repeatedly describing an approximately 15% chance of increased risk as "a small possibility" on p.102-3 and again in the response to review comments). It should also be noted that the "some increased risk" (169 deaths caused at the 95%ile) for which there is a "small possibility" is roughly twice the magnitude of an optimistic protective effect in the meta-analysis model (86 deaths averted at the 95% confidence level). Rather than attempting to use value-laden words to diminish the apparent risk presented by the pooled analysis model, the authors should explain why they have reason to believe (if in fact they should) that the pooled analysis does not represent the true range of uncertainty. As a reasonable default, a reader may be more inclined to believe that the approach with the wider confidence bounds (e.g., ranging from -1400 to +169, for a certain population group) is more believable than the one with narrower confidence bounds (-86 to -9).

II. Response to Charge Questions

1) Do you think the creation of two separate documents enhances or distracts from communication of the information contained in the report?

The creation of two separate documents is a useful change in the documentation of the assessment. The delineation, in part, reflects the distinct tasks of evaluating and interpreting the

evidence base and that of structuring and preparing a specific conceptual and computational model and analyzing its results. It is important for the reader to understand how much of the assessment process is created by the assessment team themselves, rather than the more superficial process of drawing conclusions from a literature review. While this would presumably be well beyond their current intended scope of effort, I might even suggest that future documents be separated into three documents, with the third document describing the formal nature of the calculations performed and providing further model analysis elements such as sensitivity analysis to key assumptions, importance analysis and other more analytic results. This would have a different intended audience, could be written with much less explanatory text regarding the mathematics, and would meet the needs of a more complete peer review of the effort.

2) Do you have any comments on the revisions made to the draft assessment?

The authors have made some useful steps to describe the overall approach and the underlying uncertainties (e.g., the link between the overall diagrams of Figures IV-2 and IV-3 with the knowledge gaps in Tables IV-2 and IV-3). This is very helpful for placing the overall approach and its underlying evidence base into a more complete mental model. If anything, that section should be expanded with links not only to the following tables, but to the relevant sections of the report (using the roadmap concept suggested by one reviewer). It would be useful to use this as an introduction to a more narrative treatment of the overall uncertainty in the model, putting some perspectives on uncertainties that are explicitly quantified in the model and those which are not quantified (and are thus not represented in the confidence intervals). It would also be useful to help the reader to interpret various relatively subtle outcomes of the assessment process (the additional uncertainty added by the process of generating a composite z-score across multiple measurement endpoints). Again, the micro-scale discussion (of individual knowledge gaps) is disproportionate to the need for macro-scale discussion (what level of confidence is there in some subset of the conclusions), assuming that applying an appropriate level of uncertainty in the results is the reader's primary interest.

3) Do you have any additional comments on this draft assessment?

The responses to the peer review comments appear to be quite uneven. While many reviewer comments were answered in an almost line-by-line fashion, others were amalgamated into a single concern (e.g., all of one reviewer's comments were addressed as one item: "sufficiency of documentation").

As an example of imbalance, one review comment enquired about a single data point (e.g., an inaccuracy in the central estimates for age at first walking), and was provided with a specific response to that concern. An issue that addresses all of the results presented (such as the adequacy of the simulation size to generate numerically stable central and percentile estimates) was not addressed at all. A simple audit of the impact of the number of uncertainty and variability iterations on a selection of the report's numerical estimates, and some indication in the updated report, would have been sufficient to indicate that this is not an important consideration. Ignoring the concern (or categorizing it inappropriately as merely a matter of documentation) can only raise suspicion that there is an underlying problem.

III. Specific Observations

Table V-15 (p. 106): I assume this should say "A decrease of 88 (187, -627) deaths per year." (removal of negative sign before 187 to reflect uncertainty in pooled analysis model). This is important given the prediction includes the potential for the consumption scenario to result in 187 deaths at the 95%ile.

(Below Table AA-15 on p. 155): Presumably, the note is meant to refer to Table AA-14 (rather than Table AA-1).

(p. 167): There is a reference to something called AA-19. This appears to be a result of the re-organization of the document. Presumably, the authors had meant to refer to the figures that are now Figures IV-2 and IV-3.

The few minutes that it would have required to include a table of contents in these documents was sorely missed. At the very least, this would have been very helpful in comparing the revised overall structure of the new documents, and comparing them to the previously reviewed document. It would also have been helpful to focus attention to specific parts of the new documents given the more limited effort that is allocated to the second phase of this review.

FDA Responses to the Peer Review Comments

The following responses are to peer review comments that requested changes or clarifications. The responses appear below the concerns as follows:

Concern
>   **Response**

Requested Additions:

1. Add a table of contents and a glossary.
>   We have done so.
2. Add an appendix.
>   We decline to do so due to practical difficulties although we agree that an appendix would be a desirable addition. As an alternative requested by one of the peer reviewers, we did add a "roadmap" to the assessment that fulfills some of the functions of an appendix. See request 4, below. We will consider adding an appendix at a later date.
3. Add a list of additional research needs to be targeted to critical missing data throughout the report.
>   We have done so.
4. Add a one paragraph abstract that could be disseminated as part of a health promotion by health care providers, stressing the positive benefits of eating low methylmercury-containing fish.
>   The closest we have come to a one paragraph abstract in the assessment is an executive summary that conveys the overall results, although not in quantitative terms, as well as the scientific context for those results. The assessment is not intended to impart advice to consumers on eating fish or on how much fish to eat. The draft updated consumption advice issued jointly by FDA and EPA on **[INSERT DATE]** is designed to serve that purpose. In addition, the Dietary Guidelines for Americans 2010, published jointly by the Departments of Health and Human Services, contains text in Chapter 4 that is also on point. .
5. Add a few key summary tables and figures to the Executive Summary.
>   We have kept the Executive Summary brief by design. It does contain a short summary of the key modeling results but no figures or tables. We have added an appendix (Appendix B), however, that is devoted more fully to the quantitative results and has tables that summarize the results.
6. Add a "roadmap" to the assessment.
>   We have done so. The "roadmap" provides a description of each section of text in greater detail than is provided in the table of contents.
7. Add a conclusion section (a new Section IV). Alternate request: Add a summary section before Appendix A.
>   We have added a summary appendix (Appendix B).
8. FDA should not have deleted its discussion of how some people might be adversely affected due to genetic susceptibility. This discussion should be added back, expanded, and explained more fully.

Genetic susceptibility is an area that is beginning to receive increased attention in the scientific community. Because an understanding of genetic susceptibility remains in its early stages, it is, for the most part, an area of uncertainty in our modeling. For that reason we include it as a research need in Appendix E.

Editorial Requests:

1. Fix typographical errors in the Executive Summary.

We apologize for any typographical errors that may have occurred in the draft assessment. Hopefully we have caught them all this time around.

2. The report reads much better than the earlier draft but it would still benefit from substantial editing to improve the clarity and the presentation.

The text has been revised extensively from the draft assessment. The subject is complex and presenting it clearly is a challenge. Hopefully we have done better this time.

3. Address the lack of consistency in the terms used to refer to the risk and benefit assessment itself.

We agree. We appreciate that people have had strong opinions about whether the assessment should be characterized as "risk/benefit" or "benefit/risk" or even "risk/risk." There are merits to each. Our thinking has evolved and we have renamed the assessment in order to highlight that it is a quantitative assessment of the net effects from eating commercial fish during pregnancy. The assessment now primarily refers to itself as "the assessment of net effects" or just "the assessment." We note that the concept of "net effects" is gaining international acceptance and is appearing with greater frequency in the scientific literature.

4. The document is difficult to follow, in part because it is difficult to impart a cogent argument in a lengthy document. The main text should be in proportion to that which is essential to the bottom line and the rest should be left to appendices (tactical streamlining).

We agree in principle that shorter is often better and we have, in fact, moved a significant amount of material from the main text into appendices. The drafting challenge presented by the suggested division of materials would be how to break up the text between the main body and the appendices without unduly breaking up the flow of ideas and concepts. The "bottom line" referred to by the peer reviewer is complex and includes results from a number of different modelings. In the current division, the main text is intended to be reasonably accessible to the lay reader and to provide both context and modeling results while the Appendices C and D are more technical. We prefer to leave it that way for now, with the option of drafting a more abbreviated main text at some point if need be. For now, we have: (a) focused the scope by removing the assessment of net effects on coronary heart disease and stroke and leaving it as a draft, at least for the time being; (b) rewritten the assessment substantially for clarity and other reasons; (c) moved a discussion of the modeling for early and later age verbal development into an Appendix (Appendix A) since these are essentially secondary modelings; and (d) included a summary appendix (Appendix B) for those who prefer to focus on "bottom lines."

5. Move long tables into the appendices. They interrupt the flow of the argument.

The question raised here is whether the longer tables should be separated from the text to which they relate. We agree that the long flow diagram/table combinations in Section IV can interrupt the flow of the argument so we have repositioned them in that Section in order to minimize that problem. For most tables and figures, we prefer to keep them

adjacent to the text that explains them. Otherwise, the explanatory text would have to be lengthened for the benefit of those who choose not to track down tables and figures located elsewhere. In Section V, for example, much of the text serves a guide on how to read the tables that contain the modeling results.

6. There was a reference to AA-19 in the second to last paragraph in Appendix A. Presumably the reference should have been to figures that are now figures IV-2 and IV-3.

> The reference was incorrect and has been fixed.

7. In Section V, include ounces or servings in addition to grams in the tables. Grams are not generally used by American consumers.

> We no longer use metric measurements in the main text (Sections I-V), although we continue to do so in the technical appendices (Appendices C and D).

8. The table that summarizes the results of the "what if" modeling for fetal neurodevelopment should consistently use "IQse", not "IQ point," since they are not equivalent.

> That table has been revised in a number of respects, including accommodating this comment.

Scope of the Assessment:

1. Add a statement such as "This study does not address the potential neurodevelopmental effects of ethylmercury and the relationship of mercury exposure to autism." Many consumers do not understand that this report does not address neurodevelopmental effects from ethylmercury.

> We address this point by stating at the beginning of the assessment that:
>> (a) The purpose of the assessment is to estimate the effects of maternal fish consumption on neurodevelopment;
>> (b) One aspect of doing so involves estimating the contribution made by methylmercury to those effects; and
>> (c) Methylmercury is essentially the only form of mercury in fish and thus it is the only form relevant to this assessment.
>
> We hope that this clarification is sufficient.

2. It is unclear why mercury other than methylmercury would be excluded from this analysis. If the information on inorganic mercury is available from NHANES and its effects are believe to be similar to methylmercury, why not use it?

> See the answer to Request 1, above. The quantitative assessment of net effects was developed in furtherance of FDA's regulatory responsibility for human food safety of commercial fish. We are not aware of any other form of mercury in commercial fish in amounts that are significant nor are we aware of any other form of mercury that represents a significant food safety issue.

3. The paper needs to address – at least as a limitation – that other contaminants could contribute to the net effect. Also, it should be made clear that the "what if" modeling does not include health effects from contaminants other than methylmercury.

> We agree and address the point in Section I as a "limitation" in the assessment. This limitation applies to all aspects of the assessment, including the hypothetical modeling. The evidence for adverse effects from other chemical contaminants in commercial fish is not as strong as it is for methylmercury. There are little or no data currently available to support a quantitative analysis of the possible dose-response relationships.

4. The section on postnatal exposure should not have been deleted.

A quantitative assessment of the net effects on neurological function in the general population does not appear feasible at this time. Partly for that reason, we deleted some but not all of the discussion of research relating to postnatal exposure and we are saving it for a possible future assessment. We have retained the discussion of postnatal exposure in children, however, since that matter is addressed in FDA/EPA consumption advice. We agree that it is important.

5. Section III should state that the scope is limited to consequences *in utero* rather than just to the developing nervous system, because the nervous system continues to develop after birth.
Section III has been drafted to clarify this point.

Using Delays in First Walking and Talking to Develop a Dose-Response Function for Methylmercury and Neurodevelopment:

1. Are there data to show that age of first walking and talking follow a normal distribution?
The World Health Organization (2006) multicenter study of motor development and the milestone data we have from the Seychelles Islands indicate distributions that are slightly skewed with a greater preponderance of values above the mean than below it. However, since the standard deviations are only used as scaling factors to compare estimates, we do not believe that it is critical that the underlying distribution be perfectly normal.

2. The statement that "milestone standard deviations do not vary greatly among populations" is not referenced and I do not know its accuracy for age of talking.
We have modified the sentence to read "milestone standard deviations for age of first walking do not vary greatly among populations" and we have provided a reference. As far as we know, there has not been a cross-cultural comparison of age of first talking. We have include din the modeling an uncertainty of plus or minus 10 percent to cover this point.

3. Converting ages of milestone attainment into "IQse" is questionable from a clinical and methodological perspective.
The underlying conversion here was from delays in reaching certain developmental milestones into Z-Scores. The purpose of the conversion was to enable a comparison between the size of one effect, i.e., the adverse effect of methylmercury on early milestone attainment, and the size of the beneficial effect of fish nutrients on early verbal development. The purposes of the size comparison were to: (a) determine whether and under what circumstances adverse effects are bigger than beneficial effects and vice versa; and (b) obtain some understanding of the magnitude of these effects. Using Z-Scores for purposes of comparison is not novel. A Z-Score is a tool for converting an absolute measure to a relative measure and is often used to compare different measures of psychological performance. It does not make a measure any more or less clinically relevant than it was before the conversion.

After converting time delays into Z-Scores, we then converted the Z-Scores into "IQ Size Equivalents" (a term we created) in response to a request from another peer review that we convert Z-Scores into units of measurement more understandable for most readers. We chose IQ points because there is a close relationship between Z-Scores and IQ points (IQ points are Z-Scores x 15). We named these units "IQ Size Equivalents" because they represent changes in neurodevelopment that are equivalent in size to a certain number of

IQ points. We do not claim that a delay or a loss of function in early age verbal development means a loss of IQ.

4. Age of walking and talking are crude measures and do not evaluate a whole range of more complicated developmental skills… IQ provides a better measure and is much stronger information.

        We have adopted this comment by adding IQ to the assessment as our primary modeling. On the matter of ages of first talking and walking, we continue retain the modeling that contains the effect of methylmecury on age of first talking of from the 2009 draft (early age verbal development). The results from that modeling are not identical to those for IQ but they are parallel IQ in many respects. We continue to believe that they have value, while acknowledging that there are uncertainties.

5. Using statistics to show delay in terms of hours is beyond the accuracy of the data on which it is based. Delays in hours should be eliminated from the tables. The unit of months is the highest precision that the Iraq data can support, not days or hours.

        This comment refers to the estimates of methylmercury's effect on age of first talking and age of first walking. We no longer convey results in terms of delays in hours but we do convey results in terms of days. If one month were the smallest unit of measurement that could be applied to these results, all delays would either have to be expressed as a delay of "a month" without any differentiation, or expressed as a delay of "less than a month" without stating how much less. The former would not be accurate; the latter would be accurate but not particularly informative. In considering whether the data can support calculations of specific fractions of months, the dose-response function from which these calculations are derived turns out to be very close to the dose-response function developed for IQ deficits. Given this consistency, we feel comfortable providing fractions of months, expressed in terms of days, calculated by the model.

6. "Age of first talking" needs to be clearly defined.

        We agree and have done so in Section IV and Appendix A.

7. Is age of first talking comparable with the Denver Developmental Screening Test for modeling purposes?

        Although age of first talking is not identical to a standardized test conducted at about the same age, it is not a scientific "given" that the two need to be precisely the same for purposes of this type of assessment so long as they are within the same general neurodevelopmental domain and measured at essentially the same age. We acknowledge the uncertainty, however. As was recommended to us, we have added IQ to the assessment so that we are not totally reliant on results from early age verbal development. The results from the two models are fundamentally consistent, although the IQ modeling produced somewhat greater net benefits than were estimated for early age verbal development.

8. 10.4 months is a young median age for walking. Estimates reported in the literature are about 12 months.

        The estimate is partly based on data from the Seychelles Islands, where age of first walking is lower than the world average. We do not believe that this is a problem for the modeling however, because it is the extent of change from a baseline that is important, rather than the baseline itself.

Exposure-Related Issues:

1. Differentiate between methylmercury from commercial and non-commercial fish. Use a relative source contribution model such as the lead IEUBK models to provide relative source contributions for commercial and non-commercial sources of fish.

> We would like to do this but it is currently not possible since fish consumption surveys do not typically distinguish between consumption of commercial and non-commercial fish. We do not believe that there is a basis for estimating relative contributions. However, we have increased our estimate of fish consumption by an uncertainty range of 5 – 15 percent to account for consumption of non-commercial fish. We have also added this uncertainty as a research need in appendix F.

2. In order to be able to state that exposure to methylmercury from molluscan shellfish is not of public health significance, FDA should take into account the amount eaten as well as the amount of methylmercury in the shellfish.

> The assessment does not state that exposure to methylmercury from molluscan shellfish is not of public health significance; nonetheless, we agree in principle with the comment. The amounts of molluscan shellfish eaten and the amounts of methylmercury in molluscan shellfish are included in the model.

3. Since the food recall survey used in the exposure assessment was conducted before FDA and EPA issued the 2004 fish consumption advice, women of childbearing would have decreased their fish intake and this decrease could have caused an underestimation of methylmercury exposure.

> If women of childbearing age have decreased their fish consumption since the survey, the consequence would likely be some overestimation of exposure to both methylmercury and beneficial fish nutrients. We acknowledge such as an uncertainty. It is worth noting that we have added a sensitivity analysis to the modeling to determine how an underestimation of exposure to methylmercury could affect the estimates in the assessment. In that sensitivity analysis we increased the amount of methylmercury in all fish consumed by 20 percent.

4. Modeled hair levels of mercury are between 30-40 percent lower than were reported in the NHANES survey. Two possible explanations are provided and while there is some plausibility to them, another possible explanation is that the exposure modeling is systematically underestimating exposures. There also does not seem to be any reflection of this possibility in the confidence intervals.

> We have revised the hair-blood model so that our estimates are closer to the NHANES estimate with overlapping confidence intervals. The sensitivity analysis that we have added, as described in the previous response, effectively examines how the assessment results would differ if the exposure assessment produced results very close to those that have been reported by NHANES.

Beneficial "Fish" Contribution to Net Effects:

Some greater context [additional data from other studies] was needed for results on this topic.

> We agree. It is one of the reasons why we added modeling on the net effects of fish consumption during pregnancy on IQ. In the 2009 draft, the data that we modeled on the beneficial effects of fish consumption were limited to results on early age tests of verbal development. As a caveat, however, the two studies come from the United

Kingdom and there is some overlap in the study populations. We regard that as a limitation that will have to be addressed in future modeling.

Individual Fish:

1. The document should acknowledge that for regular consumption of certain kinds of fish, methylmercury could present a problem.

> We agree. This is an important point.   To address it, we have modeled the net effects on fetal neurodevelopment that could occur from eating individual species of fish. For 47 species and market types of commercial fish, we have estimated three data points: (1) how much fish would have to be eaten per week in order to obtain the maximum beneficial net effect, if any, that can be conveyed by that particular species or market type; (2) the size of that maximum beneficial effect; and (3) the amount of fish that would have to be eaten per week in order for the net effect to be adverse.

2. The table that summarizes mercury concentrations in fish contains no information about shrimp or canned salmon.

> The mercury concentrations in shrimp and salmon are included in all the tables in the assessment that provide information on concentrations (there are several such tables). Given the popularity of shrimp and salmon, we would be remiss if we did not include those fish.

Various Uncertainties in the Model:

1. The executive summary should address the issue of underestimation of quantitative dose-response relationships derived from questionnaire estimates vs. biomarkers.

> The executive summary is a brief overview of the assessment that, by design, does not contain that level of detail. However, we have added this point in Table IV-2, which specifically addresses various uncertainties and assumptions in the modeling.

2. There should be more discussion of uncertainty at the level of results, as compared to the uncertainty in any single element of the model. The micro-scale discussion (of individual knowledge gaps) is disproportionate to the need for macro-scale discussion (what level of confidence is there in some subset of the conclusions), assuming that applying an appropriate level of uncertainty in the results is in the reader's primary interest.

> All of the uncertainties that we regarded as potentially significant were accounted for in the model and are thus reflected in the confidence intervals. From a qualitative standpoint, in addition to tables that address the significance of the many of the uncertainties, we have included text that points out the various uncertainties as reflected by the confidence intervals when to these uncertainties have implications that are potentially important. For example, the text point out when a central estimate shows a beneficial result but one of the confidence limits indicates the possibility of an adverse result, or when a central estimate is adverse but the confidence interval includes zero. Beyond these things, we are not sure what more we could have said regarding uncertainties other than to review them in a lengthy tables (IV-1 and 2), as we do.

3. The link between the overall diagrams of Figures IV-2 and IV-3 [flow diagrams of the exposure modeling and the dose response modeling, respectively] with the knowledge gaps in Tables IV-2 and IV-3 [now Tables IV-1 and IV-2] is helpful for placing the overall approach and its underlying evidence base into a more complete mental model. The section should be expanded with links to following tables and to relevant sections of the report (using the roadmap

concept suggested by one reviewer). It would be useful to use this as an introduction to a more narrative treatment of the overall uncertainty in the model.

> In each row in Tables IV-1 and IV-2, we have added cites to relevant sections of the assessment to accommodate this comment. Beyond that, we are not sure what more we could do on the issue of uncertainties, as discussed above.

4. There is need for a detailed (and independent) check of the models and data in addition to peer review.   I agree with the selection of the data set, models, and the apparent application of the models to the data…but I have largely assumed that the software has actually done the calculation as described. There need to be independent confirmation that the actual calculations have occurred as described (by independent, I would suggest that another FDA staff member randomly check different components of the model).

> We obtained an independent review by a modeling expert unconnected with this assessment to determine whether the model actually does the calculations as described in the report. This review focused on the codes developed for the assessment, the simulation aspects of the model, and reproducing the major parts of the model (the exposure assessment, the dose-response and the risk characterization) in another platform. The procedure required that each line of modeling code was read, understood and validated. The process identified some errors that were immediately corrected. After correction, the reviewer advised us that that the model is robust and that the conclusions of the assessment are supported.