Peer Review Report

Peer Review of
Idaho DEQ Fish Consumption Survey

January 22, 2016

Peer Reviewers:

Patricia M. Guenther, Ph.D., RD
Alanna J. Moshfegh, MS, RD
Janet A. Tooze, Ph.D., M.P.H.

Idaho DEQ Contract #K109

Prepared for:
Idaho Department of Environmental Quality
1410 N Hilton
Boise, Idaho 83706

Prepared by:
Versar, Inc.
6850 Versar Center
Springfield, VA 22151

TABLE OF CONTENTS

I. INTRODUCTION

To characterize fish consumption rates for Idaho residents, Idaho Department of Environmental Quality (DEQ) undertook a survey of residents on their fish eating habits. Responses from the survey questions were analyzed to determine rates of fish consumption, ultimately a daily rate for each respondent. Information was collected regarding fish consumption rates for two populations in Idaho: (1) the general population, and (2) recreational anglers. The objective was to determine the distributions of fish consumption rates for these Idaho populations, in addition to traditional point estimates, such as a 90th percentile rate. The approach, methodology, and findings from this effort are summarized in the draft document "Idaho Fish Consumption Survey," authored by Northwest Research Group (NWRG). These data will be used, in conjunction with NCI analysis of the fish consumption survey data, by the Idaho DEQ in deriving water quality criteria to be protective of human health. Additional background on the larger scope of Idaho DEQ's efforts can be found in the on Idaho DEQ's web site: http://www.deq.idaho.gov/58-0102-1201.

The purpose of this peer review was to have three experts evaluate the draft document "Idaho Fish Consumption Survey" with regards to the survey design and implementation, including the calculation of fish consumption rates for all adult Idahoans and adult Idahoans who are anglers for fish of all types and source (All Fish) and for resident freshwater fish (Idaho fish).

Peer Reviewers:

Patricia M. Guenther, Ph.D., RD
Guenther Consulting
Salt Lake City, UT

Alanna J. Moshfegh, MS, RD
Agricultural Research Service, USDA
Beltsville, MD

Janet A. Tooze, Ph.D., M.P.H.
Wake Forest School of Medicine
Winston-Salem, NC

II. PEER REVIEW PROCESS

This section details the procedures that were followed to conduct this external peer review. Versar has a well-established approach for conducting peer reviews, from completion of more than 500 peer reviews of environmental risk assessment-related documents over the past 20 years for a variety of Federal and state government research and regulatory agencies. The approach covers all aspects of the peer review, from reviewer selection through completion of the peer review report. Within this approach are several quality assurance protocols to ensure that: qualified individuals are selected to participate, they are free from conflict of interest (COI) and the appearance of the lack of impartiality, and a thorough review is completed.

Reviewer Identification and Selection

Versar's approach for selecting the technical expert reviewers consisted of the following five key steps: (1) development of selection criteria, (2) identification of experts, (3) COI screening, (4) selection of peer reviewers, and (5) confirmation of peer reviewer participation.

The experts that participated in this review were identified by literature searches of scientific journals, professional societies, and scientific meetings, as well as searches of Versar's internal peer review database of more than 3,000 scientists. As a result of this search, Versar identified potential scientific experts with expertise in the general area fish consumption surveys. These experts were contacted to determine their availability and interest in participating in the review. Interested candidates provided their *curriculum vitae*, which were reviewed by Versar staff to ensure that each candidate had the appropriate scientific credentials and evidence of expertise through a listing of their publications and professional affiliations. The specific disciplines/areas of expertise needed for this peer review included: (1) fish consumption survey design and implementation, (2) computation of food consumption statistics from survey data collected using a food frequency questionnaire (FFQ) approach, and (3) statistical modeling of short term nutritional survey data to produce fish consumption statistics using the National Cancer Institute (NCI) method.

Versar also conducted COI screening to make certain that the experts had no COI or appearance of the lack of impartiality that would interfere with providing a thorough critical review of the document. This screening involved sending the potential candidates a series of COI screening questions that helped us to determine if they were involved with any other work and/or organizations that might create a real or perceived conflict of interest for the current task. Additionally, each expert signed forms certifying that, to the best of their knowledge, they did not have any conflict of interest related to the task. Upon completion of the COI screening, Versar selected three experts, based on their credentials, to conduct the review. Versar requested and received consent from the Idaho DEQ and, subsequently, contacted the three reviewers to notify them that they were selected to participate in the peer review.

Conducting the Review

Following the selection process, Versar distributed to the reviewers the draft document "Idaho Fish Consumption Survey" and a work assignment authorization (WAA) letter, which included the charge questions, instructions, and a comment template for the preparation of written comments to ensure that each reviewer submitted their comments in a consistent format.

Versar developed a series of charge questions to help guide and focus the review of the document. These charge questions generally asked reviewers to provide comments on the strengths, weaknesses, and overall quality of the study and report. The comment template instructed the reviewers to provide comments in the following three categories:

(1) General Impressions - overall comments addressing the accuracy of information presented, clarity of presentation, and soundness of conclusions.

(2) Response to Charge Questions - narrative responses to the five charge questions.

(3) Specific Observations - Specific observations or comments on the document, including editorial changes, mentioning page and line.

The WAA also included Versar's confidentiality statement indicating that the peer reviewers' should not distribute or discuss their comments with any outside party, as well as the amount of time the external reviewers had to complete their reviews and submit written comments. During the review period, Versar monitored the progress of the reviewers on a regular basis to make sure there was timely delivery of the written comments.

Review of Expert Comments

At the completion of the review period, Versar received written comments from the three reviewers, evaluated the experts' comments for completeness and scientific quality, and obtained clarification or additional input as needed. The three experts all submitted thorough reviews of the draft report, providing: (1) general comments, which included their overall impressions of the document, addressing the accuracy of information presented, clarity of presentation, and soundness of the conclusions; (2) responses to five charge questions; and (3) specific observations, which included editorial corrections or factual changes to the document. The comments were compiled into a peer review summary report, and organized by charge question to facilitate side-by-side viewing of the reviewers' comments on the same topics.

III. CHARGE TO REVIEWERS

Charge Questions:

1. Please comment on the clarity and organization of the report. Does it present information, including tables and figures, in a clear and usable format? If not, please provide suggestions for improving the clarity of the document, which is intended to be useful to state regulators, the scientific community, and other stakeholders, including the general public.
2. Please comment on the appropriateness of the survey design and implementation methodologies, including the food frequency questions and dietary recall questions, for the intended purpose?
3. Please comment on the analysis and findings for differences in the portion size estimates between food frequency and dietary recall questions.
4. Are the results of the fish consumption survey scientifically sound and "valid" for use, in conjunction with the NCI method, for derivation of water quality criteria to be protective of human health for the general population and recreational anglers?
5. Do you have any other suggestions for improving the scientific quality or utility of the document?

IV. GENERAL IMPRESSIONS

Patricia M. Guenther

The report is poorly organized, making it very difficult to comprehend. Much of the descriptive statements about the design of the survey and about the questions seems to be inaccurate (examples are listed below), and much information is not clearly presented. The purpose of the report is not clearly stated, and the results are not presented in the context of the purposes. Quite a bit of seemingly extraneous information is provided. A lot of methodology work was done prior to fielding the survey, but it is mixed in with survey results when it should be reported separately, perhaps in an appendix. A conclusion or summary of highlights or results section is missing.

Alanna J. Moshfegh

This report provides a comprehensive and detailed description of design, execution, and findings of a research study on fish consumption of adults 18 years of age and older that reside in Idaho. Anglers were specifically over-sampled (33% of sample) in the study design. The information presented is based on a representative fish consumption survey of this sample. The information presented is extensive with thoroughness in the details of the sample design, data collection methodologies design, testing, and execution, and analyses of that data. Each step and consideration in the sample design and data collection execution are described giving the reader a thorough understanding and appreciation that nothing was left to chance and the study team did not fail to address issues and limitations related to this type of study. The information presented is quite clear in presentation and conclusions reached are supported by the data.

Janet A. Tooze

In general, the methodology used to develop the surveys seemed sound, and the implementation seemed reasonable, with the possible exception of the low response rate. While these studies are generally prone to error, it appears that several steps were taken to minimize error. It appears that the survey weights were appropriately calculated from the survey. It also appears that the data were properly processed. I would have liked to have seen more discussion regarding whether these estimates represent usual dietary intake, particularly from the recalls. It is not clear if anyone with a nutrition background was involved in this process. Ideally a registered dietitian would have conducted the training and been involved in the design and interpretation of the analysis. It is concerning that the probability of reporting consumption on recalls decreases over time, and the impact of this was not addressed in a sensitivity analysis in the report. Statistical methods in general are not well described. There is no discussion of why normality was assumed with skewed data. It is not always clear if non-consumers are included or excluded in some statements. Some confidence intervals do not appear to be reasonable. In general, I would have liked more synthesis of all the information in this report, and more sensitivity analysis exploring differences between methods and assumptions made. For example, did the type of food (mixed dish or not), impact the portion size estimation? What happens if not all days of recall were used? I am also not sure of the reason for not trying to integrate the two reports (this one and the IMS report), but I think this really would help to improve clarity of the analysis of this survey. In addition, it may be useful to have IMS analyze some of the FFQ data as well as the recalls.

V. RESPONSE TO CHARGE QUESTIONS

Charge Question 1
Please comment on the clarity and organization of the report. Does it present information, including tables and figures, in a clear and usable format? If not, please provide suggestions for improving the clarity of the document, which is intended to be useful to state regulators, the scientific community, and other stakeholders, including the general public.

Patricia M. Guenther

The report overall is not clear and is poorly organized. The results are not presented in a usable format. Methods and results are intermingled; they should be clearly separated and presented in separate sections.

When writing this report, the authors did not fully consider the basic principles of survey research: a sample of the target population is selected; data from the sample is collected; the data is then weighted to account for the survey design and statistically analyzed to provide estimated parameters of interest for the target population. Results should be referred to as estimated population parameters (for example, estimated proportions, means, medians, and percentiles) and not in terms of what the characteristics of the survey participants or in terms of what they said or did. The methods section should discuss the survey participants; the results section should pertain to the adult population of Idaho and to adult anglers in Idaho.

The authors are inconsistent about the length of the reference period for the recalls, saying in most places that it was 8 days and in several places, 7 days. The questionnaire itself clearly indicates that is was 7 days, but Figures 13 and 14 report data from 8 days. A bigger problem, however, is that it appears data were actually collected for only 1 to 2 days during each phase of the study. According to the questionnaire, if a respondent reported consuming fish on the 24-hr recall, then he/she was asked if fish was eaten on any of the 6 days prior to the recalled day, but only 1 additional day of fish intake data was collected. If a respondent did not reported fish on the 24-hr recall, then he/she was asked if fish was eaten on any of the previous 7 days, but only 1 day of fish intake data was collected. The questionnaire does not include any questions asking about additional days or a skip pattern for any such questions. Some results suggest that a maximum of 2 days were collected as well (Figures 13 and 14 and Tables 39 and 40). The data file should be checked to see if any respondent has more 2 days of fish intake reported. If only a maximum of 2 days were included, this cannot be characterized as a 7 (or 8)-day recall. A 7- or 8-day recall would include *all* the fish eaten during that period.

In order to be usable, the results on fish consumption must indicate if the weight of fish reported (i.e., ounces/grams) represents cooked, edible portion, that is, does it consider refuse, moisture loses, and fat gains (in frying), or does it represents something else, such as raw weight of whole fish including refuse? Dietary surveys, including fish-specific surveys, typically use conversion factors to get from the estimated ounces of cooked fish reported to the grams of fish in the form needed for the analysis.

Alanna J. Moshfegh

The report is quite clearly written with easy to understand, clear graphics that help the reader to quickly realize the major results. Each section is labeled with detailed sub-sections for ease in locating information and understanding details. There are some minor suggestions I offer for improving clarity or discussion which I have detailed under "Specific Observations."

Janet A. Tooze

The sections of the report are well organized, and results are presented in a logical format. However, the details as to which methods were used to test statistical significance are lacking from the tables (and even in much of the text). The tables should stand alone, and should mention the statistical methods used to determine statistical significance, and should include sample size (some do, but not all). Also the table titles should clearly indicate the source and the time period. Some tables like Table 17 have subscripts with no explanation of what they mean.

Charge Question 2
Please comment on the appropriateness of the survey design and implementation methodologies, including the food frequency questions and dietary recall questions, for the intended purpose?

Patricia M. Guenther

This is difficult to assess because "the intended purpose" of the study is unclear. On page 12, it says, "…it desirable to know the distribution of fish consumptions rates within the population surveyed. … IDEQ expects that ultimately it will be necessary to be able to compare expected risks for the general population to expected risks for recreational anglers, members of Indian tribes in Idaho, or others with higher consumption than the general population. In addition, IDEQ also wants to be able to say what proportion of the overall fish consumption rate (i.e., from all sources) represents fish caught from Idaho waters, and within the latter what proportion of fish species is resident versus migratory." So the reader might assume this is the purpose of this report. However, other extraneous information is provided, and no reason is given for including it. Resident versus migratory fish is not addressed; a reason for this should be provided.

A clear and concise statement of purpose of this study is needed. This is important because seemingly extraneous information was collected (to the detriment of overall quality of the study). The report mentions some additional things that are "desirable" (p. 11), but it is unclear whether or not these were actually part of the purpose.

Alanna J. Moshfegh

The most current methodology for food frequency and 24-hour dietary recall were used and aptly adapted for collection of fish consumption only in this study. It is commendable that in addition to a 24-hour dietary recall, the contractor added a 7-day food record to ascertain intake of fish beyond just one day but up to 8 days. Details about pilot testing, interviewer training, and monitoring of data collection on 10% of the sample reflect a strong survey operation.

Janet A. Tooze

The methodologies used were appropriate for the intended purpose of measuring dietary intake. Research has shown that food frequency questionnaires (FFQs) are subject to substantial measurement error, including both random and systematic error (bias). This is because they are quite cognitively challenging, requiring people to do averaging in their heads over a fairly long time period. One strength of this study was the cognitive testing and pilot testing of the questionnaires. These procedures would be expected to minimize errors with clarity of the questions and consistency of administration. All surveys were administered by phone, which is generally considered to yield similar results to in-person administration, and the pilot testing done by NWRG supported this as well, so that seems to be a valid way to administer the surveys. The cognitive testing supported asking participants to recall fish consumption by location (home, restaurant, other) and by meal, which hopefully minimized errors in recall. Because the FFQ was only focused on fish consumption, it also queried a large number of mixed dishes that contained fish, which should lead to fewer assumptions that need to be made regarding recipe content, which is also a strength. Participants were not asked about seasonal fish consumption or fish consumption at special events or to specify amounts for different types of fish, which might have improved estimation of intakes.

An 8-day recall was used to assess fish intake as well. This method is not commonly used, but would be expected to have errors similar to a 24-hour recall, which is more commonly used to assess short term intake. In general, short-term recalls are also prone to both random and systematic error like FFQs, but recalls tend to have less systematic error than FFQs.

Furthermore, statistical methods such as the NCI method may be used to obtain estimates of the distribution of usual intake that are adjusted for random error. Figure 13 is of concern, that recall of fish intake went down as time from the interview increased. This suggests a systematic error in the 8-d recall that would be expected to lead to an underestimate of true fish consumption, as one would expect the 'yesterday' recall to be the most accurate. It would be helpful to do some analyses where intake is normalized to the first day and sensitivity analyses of this issue.

Charge Question 3
Please comment on the analysis and findings for differences in the portion size estimates between food frequency and dietary recall questions.

Patricia M. Guenther

Different dietary data collection methods should be expected to yield different results.

Generally speaking, when a survey respondent is asked the same thing twice, the second answer has a value that is less than the first; and, generally speaking, the first answer is the more accurate. This rule of thumb held in this study when respondents reported smaller portion sizes of fish during the second recall period than in the first. The conclusion that the portion sizes reported for the first recall period are more accurate than in the second follows this paradigm. It is also reasonable to assume that respondents recalled the number of fish-containing snacks,

which would be expected to have smaller portion sizes, more accurately during the recalls than on the FFQ, as the authors suggest.

Alanna J. Moshfegh

The differences found between food frequency and dietary recall on fish consumption is not unlike other studies. Food frequency generally overestimates number of times an item is consumed and is also less reliable on portion size estimation of compared to 24-hour recalls. Recalls of very limited number of days may underestimate times an item is consumed, especially for those foods not typically consumed daily, such as fish. Thus, the follow-up recontact survey provides critical data to allow for usual intake analysis of fish consumption.

Janet A. Tooze

I think that the reasons for the differences between the two methods include:

(1) The averaging component that is required for the FFQ. Participants must average across the past 12 months and across mixed dishes and non-mixed fish consumption. This is quite challenging, and might lead to smaller portion sizes to balance mixed dishes from non-mixed.

(2) The recall asks for consumption of a specific meal and type of fish, which does not require averaging and would be expected to be more accurate.

(3) The FFQ asked about "number of events" and the portion size of each event (meal), whereas the recalls asked about the "number of helpings" at each event and the portion size of each helping. This might result in larger portion size estimation as well.

(4) One point I wanted clarification on is whether it is true that all the portion size questions where they were asked to provide a range coded both "between 13 and 16 oz" and "more than 16 oz or more than one lb" as "06" (see for example FFQ5A_1 and FCR24_7B_1, but it seems to be true for all frequency questions)? Of course if both were given the same code in the database this could lead to inaccuracies for the FFQ and recall.

(5) The means in the report are for skewed data, and therefore the mean is not the best measure of central tendency of the distribution.

(6) There is no measurement error adjustment. It would be interesting to compare to the consumption day amounts from the NCI method with these estimates, but this does not appear to have been done.

Charge Question 4
Are the results of the fish consumption survey scientifically sound and "valid" for use, in conjunction with the NCI method, for derivation of water quality criteria to be protective of human health for the general population and recreational anglers?

Patricia M. Guenther

It is not possible to determine if the data produced by the survey are fit for use because the report has too many problems. If they were to be addressed, and if the report could be whittled down to describe only how the survey was designed; how the data were collected, processed, and weighted; and if it would state the limitations of the design, data collection and processing methods, and weighting procedures, then perhaps it could be concluded that the data are fit for use. The authors of the companion report, which implements the NCI method, believe they were provided with recalls that have 8 days of intake data in each recall period, but that does not appear to be true. Therefore, the results presented in that report are probably not scientifically sound or fit for use.

Alanna J. Moshfegh

Yes, the results are sound and valid for use.

Janet A. Tooze

There are limitations of these data. First, they rely on self-report and methods that are known to be subject to error. Second, the response rate was low, with only 25% of potential participants participating. The participation rate of 40% was better, but there is still potential for non-response bias. There is no adjustment for measurement error in this report. However, techniques were implemented to reduce potential bias in the surveys. Overall, I would conclude that the methodology used to implement this survey was scientifically sound, and the results are informative. I would rely more on the NCI method results and the recall results (particularly the first day), rather than the FFQ results to provide "valid" estimates.

Charge Question 5
Do you have any other suggestions for improving the scientific quality or utility of the document?

Patricia M. Guenther

Other suggestions are offered below in Specific Comments.

Alanna J. Moshfegh

Specific comments provided below.

Janet A. Tooze

As I mentioned previously, I think it is important to describe the specific statistical tests that were used for each analysis. This is often not clear in the document. Sometimes statements are not properly cited in the text, or inadequate details are given to back up statements that are made. I have tried to specify many of these comments in the next section.

VI. SPECIFIC OBSERVATIONS

Specific Observations and Comments on "Idaho Fish Consumption Survey"		
Page Number	**Line Number**	**Comment or Question**
3-7		Detailed Table of Contents excellent for finding information in report.
8	3	HHC should be defined here.
8	Line 4	Why not give the same type of figure found in this study somewhere in the summary? Is that figure the 31.23 gm reported in page 9, line 16? I am assuming the basis of that number is not the same as the 17.5 gm in line 4.
9	4-19	It is not clear if these are rates for consumers only, or if they include the 12% non-consumers. This is also true in the results section.
9	11 & 29	Why not report same amount of fish by food frequency collection in these two places?
10	3-5	This statement needs to be substantiated. How is it clear?
10	15-17	How was it assessed to be "significantly lower"? Did you take into account one asks about 1 year and one asks about 8 days? Did you compare the question on the FFQ that asked about the past 7 days to the recall? I'm afraid a statement like this is just going to confuse people – of course we would expect less consumption for a 1 week period compared to a 1 year period.
10	19-20	Why is this true? Is there a reason you would assume this?
11	Line 7	First sentence does not make sense. Something is missing.
11	28	The objective I believe is to assess usual dietary intake of fish. I think calling it a "daily rate" is not quite right – it is a usual rate expressed on a daily basis, but as a long run average.
14	10-12	Please state what these strengths and weaknesses are.
21	17	I do not think it is accurate to say that FFQs are representative of usual intake. This is not cited here.
21	18	I do not think FFQs are a preferable method for foods with day to day variability. It does not state to what they are preferred, and there is no citation for this.
21	19	I think this statement is debatable. Even highly literate subjects have trouble estimating year-long averages over multiple types of fish in their head. This is not cited.
23	5-7	Not being as dependent on memory as an FFQ is actually a strength of dietary recalls, particularly for 1 day. I would not cite it as a weakness, although there is some evidence of weakness for this multi-day recall used in this study.
23	17	The research cited here was on a 7-day diary, not a 7-day recall. In a diary, the food is recorded as it is eaten. This is an inappropriate interpretation of this study.
24	39-43	In light of these findings, why was a 7-day recall period used rather than 3 or 4 days?
25	4-6	Please provide more detail about the basis/evidence for this decision.
25	5	These are locations, not meals types.

Specific Observations and Comments on "Idaho Fish Consumption Survey"		
Page Number	**Line Number**	**Comment or Question**
26	8-11	This is a very interesting study. The authors should consider publishing it in a peer-reviewed journal to disseminate the findings.
27	18, last bullet	Since income is a key variable, mention here the broad categories that were monitored. Were there only two--above and below $25,000—or were there others?
30		Excellent to have response rates by season for this research.
31		Response rates were low, but expected in a telephone survey. I did not see, but it might be helpful to include length of time of the interviews—does give appreciation to those who completed the survey.
33	14	This table seems to be more about the recontact than the main study. Please revise the title and the row headings.
35	5	Please say who recommended this approach.
44	4 and 20	"Work" should be "restaurant."
45	6	The questionnaire indicates yesterday and the 6 (not 7) days prior to yesterday.
46	9 and 10	For the recall, respondents were not asked to report "average" portion sizes. They were asked to report the size of each portion of fish eaten.
46	11	Please cite a source for the deck of cards equals 6 oz of fish. A deck of cards is commonly used to represent 3 oz of meat or poultry; however, this reviewer is unaware of the basis for that convention.
47	28-31	It is common practice to assign a default amount when the respondent cannot. This could also be imputed from the amounts that were reported. It is also common practice to assign "species, not specified." For species not specified, it would still be possible to determine if the fish was "caught by you or someone you know" and if so, if it came from Idaho waters or not.
48	7	This is incorrect. If they didn't know, they were assumed to be non-anglers.
48	19-20	Please confirm that this was true in most instances. It seems like the distribution of almost everything in this report would be skewed. Please be more specific.
48	20	Why were the assumptions made that the data were normally distributed when they appear to be so skewed?
49		Please include a 95% CI for the 88%.
49		Figure 2 results were nicely consistent with the 2012 Behavioral Risk Factor Survey.
51	Column headings	What does "n_w" mean? Please explain in a footnote.
51	Text box	How can it be that fish consumers are equally like to be male or female while fish non-consumers are more likely to be women than men? Isn't everyone either a fish consumer or not?
51		What are the subscripts? What statistics were used?
54	Text box	Please confirm that consuming 2,500 fish meals and snacks in 1 year was judged to be plausible. That's an average of 8 per day.

Specific Observations and Comments on "Idaho Fish Consumption Survey"		
Page Number	**Line Number**	**Comment or Question**
54	Text box	The median is better than what? The 99th percentile is not a measure of central tendency. The median is generally considered a better measure of central tendency than the mean for skewed data, as is the case here, but the mean is not mentioned in this paragraph.
54		This should specify a time period – annually.
54		Figure 5. Suggest giving average for fish consumers in the discussion.
56		Figure 6. Any discussion as to why number of meals with fish so much higher in October for anglers—popular/good fishing time???
56		What type of test was used?
57		Figure 7. Designation of 51-300 for heavy seems too broad a range, could ranges/further discussion be provided?
63		"Snack" is not a place. Consumption events are either meals or snacks. This analysis is inappropriate and should be omitted.
69		If someone didn't eat any fish at the specified place, their portion size was counted as 0; but if they did eat at the specified place but didn't report their portion size, they were omitted from the analysis (hence the varying n's), right? This does not make sense. Please omit this analysis.
70	Text box, para. 2	This is an incorrect interpretation of the table. Why was the 65th pctl calculated?
71		These CI seem to be way too narrow. How were they computed?
80	4-6	The evidence provided does not support this statement.
80	4-6	How is it clear?
80	17	This seems low, as 50% is the standard for most survey studies. Please cite this. This is not a public opinion study.
81	6-9	Please reconsider re-weighting the sample to fix the income distribution; that will help with the proportion of anglers, too. Is it possible to re-weight the sample to make it more representative of Idaho without worrying so about the health districts? IDEQ does not need estimates at the health district level.
99	10-25	Elsewhere in this report it says that all of the 7 days that had fish were asked about, but this set of questions only asks about the most recent day that had fish. If additional days were asked about, please add the question that asked about another day with fish and the skip pattern(s).
117	Text box, para. 1	This does not seem to be true because the questionnaire does not indicate that questions about fish consumption were asked for more than "yesterday" and one additional day. Other results suggest that this was the case as well (Figures 13 and 14 and Tables 39 and 40). Please check the data file and see any if respondent has more 2 days of fish intake reported.

VII. INDIVIDUAL REVIEWER COMMENTS

Review By:
Patricia M. Guenther, Ph.D., RD

Peer Review Comments on "Idaho Fish Consumption Survey"

Patricia M. Guenther, Ph.D., RD
Guenther Consulting
November 17, 2015

I. GENERAL IMPRESSIONS

The report is poorly organized, making it very difficult to comprehend. Much of the descriptive statements about the design of the survey and about the questions seems to be inaccurate (examples are listed below), and much information is not clearly presented. The purpose of the report is not clearly stated, and the results are not presented in the context of the purposes. Quite a bit of seemingly extraneous information is provided. A lot of methodology work was done prior to fielding the survey, but it is mixed in with survey results when it should be reported separately, perhaps in an appendix. A conclusion or summary of highlights or results section is missing.

II. RESPONSE TO CHARGE QUESTIONS

1. Please comment the clarity and organization of the report. Does it present information, including tables and figures, in a clear and usable format? If not, please provide suggestions for improving the clarity of the document, which is intended to be useful to state regulators, the scientific community, and other stakeholders, including the general public.

The report overall is not clear and is poorly organized. The results are not presented in a usable format. Methods and results are intermingled; they should be clearly separated and presented in separate sections.

When writing this report, the authors did not fully consider the basic principles of survey research: a sample of the target population is selected; data from the sample is collected; the data is then weighted to account for the survey design and statistically analyzed to provide estimated parameters of interest for the target population. Results should be referred to as estimated population parameters (for example, estimated proportions, means, medians, and percentiles) and not in terms of what the characteristics of the survey participants or in terms of what they said or did. The methods section should discuss the survey participants; the results section should pertain to the adult population of Idaho and to adult anglers in Idaho.

The authors are inconsistent about the length of the reference period for the recalls, saying in most places that it was 8 days and in several places, 7 days. The questionnaire itself clearly indicates that is was 7 days, but Figures 13 and 14 report data from 8 days. A bigger problem, however, is that it appears data were actually collected for only 1 to 2 days during each phase of the study. According to the questionnaire, if a respondent reported consuming fish on the 24-hr recall, then he/she was asked if fish was eaten on any of the 6 days prior to the recalled day, but only 1 additional day of fish intake data was collected. If a respondent did not reported fish on the 24-hr recall, then he/she was asked if fish was eaten on any of the previous 7 days, but only 1 day of fish intake data was collected. The questionnaire does not include any questions asking about additional days or a skip pattern for any such questions. Some results suggest that a

maximum of 2 days were collected as well (Figures 13 and 14 and Tables 39 and 40). The data file should be checked to see if any respondent has more 2 days of fish intake reported. If only a maximum of 2 days were included, this cannot be characterized as a 7 (or 8)-day recall. A 7- or 8-day recall would include *all* the fish eaten during that period.

In order to be usable, the results on fish consumption must indicate if the weight of fish reported (i.e., ounces/grams) represents cooked, edible portion, that is, does it consider refuse, moisture loses, and fat gains (in frying), or does it represents something else, such as raw weight of whole fish including refuse? Dietary surveys, including fish-specific surveys, typically use conversion factors to get from the estimated ounces of cooked fish reported to the grams of fish in the form needed for the analysis.

2. Please comment on the appropriateness of the survey design and implementation methodologies, including the food frequency questions and dietary recall questions, for the intended purpose?

This is difficult to assess because "the intended purpose" of the study is unclear. On page 12, it says, "…it desirable to know the distribution of fish consumptions rates within the population surveyed. … IDEQ expects that ultimately it will be necessary to be able to compare expected risks for the general population to expected risks for recreational anglers, members of Indian tribes in Idaho, or others with higher consumption than the general population. In addition, IDEQ also wants to be able to say what proportion of the overall fish consumption rate (i.e., from all sources) represents fish caught from Idaho waters, and within the latter what proportion of fish species is resident versus migratory." So the reader might assume this is the purpose of this report. However, other extraneous information is provided, and no reason is given for including it. Resident versus migratory fish is not addressed; a reason for this should be provided.

A clear and concise statement of purpose of this study is needed. This is important because seemingly extraneous information was collected (to the detriment of overall quality of the study). The report mentions some additional things that are "desirable" (p. 11), but it is unclear whether or not these were actually part of the purpose.

3. Please comment on the analysis and findings for differences in the portion size estimates between food frequency and dietary recall questions.

Different dietary data collection methods should be expected to yield different results.

Generally speaking, when a survey respondent is asked the same thing twice, the second answer has a value that is less than the first; and, generally speaking, the first answer is the more accurate. This rule of thumb held in this study when respondents reported smaller portion sizes of fish during the second recall period than in the first. The conclusion that the portion sizes reported for the first recall period are more accurate than in the second follows this paradigm. It is also reasonable to assume that respondents recalled the number of fish-containing snacks, which would be expected to have smaller portion sizes, more accurately during the recalls than on the FFQ, as the authors suggest.

4. Are the results of the fish consumption survey scientifically sound and "valid" for use, in conjunction with the NCI method, for derivation of water quality criteria to be protective of human health for the general population and recreational anglers?

It is not possible to determine if the data produced by the survey are fit for use because the report has too many problems. If they were to be addressed, and if the report could be whittled down to describe only how the survey was designed; how the data were collected, processed, and weighted; and if it would state the limitations of the design, data collection and processing methods, and weighting procedures, then perhaps it could be concluded that the data are fit for use. The authors of the companion report, which implements the NCI method, believe they were provided with recalls that have 8 days of intake data in each recall period, but that does not appear to be true. Therefore, the results presented in that report are probably not scientifically sound or fit for use.

5. Do you have any other suggestions for improving the scientific quality or utility of the document?

Other suggestions are offered below.

III. SPECIFIC OBSERVATIONS

Specific observations, comments, and questions on the "Idaho Fish Consumption Survey" document are listed below. **Note:** Additional specific comments and questions are provided in the attached pdf of the document.

Specific Observations and Comments on "Idaho Fish Consumption Survey"		
Page Number	**Line Number**	**Comment or Question**
23	17	The research cited here was on a 7-day diary, not a 7-day recall. In a diary, the food is recorded as it is eaten. This is an inappropriate interpretation of this study.
24	39-43	In light of these findings, why was a 7-day recall period used rather than 3 or 4 days?
25	4-6	Please provide more detail about the basis/evidence for this decision.
25	5	These are locations, not meals types.
26	8-11	This is a very interesting study. The authors should consider publishing it in a peer-reviewed journal to disseminate the findings.
27	18, last bullet	Since income is a key variable, mention here the broad categories that were monitored. Were there only two--above and below $25,000—or were there others?
33	14	This table seems to be more about the recontact than the main study. Please revise the title and the row headings.
35	5	Please say who recommended this approach.
44	4 and 20	"Work" should be "restaurant."
45	6	The questionnaire indicates yesterday and the 6 (not 7) days prior to yesterday.
46	9 and 10	For the recall, respondents were not asked to report "average" portion

Specific Observations and Comments on "Idaho Fish Consumption Survey"		
Page Number	**Line Number**	**Comment or Question**
		sizes. They were asked to report the size of each portion of fish eaten.
46	11	Please cite a source for the deck of cards equals 6 oz of fish. A deck of cards is commonly used to represent 3 oz of meat or poultry; however, this reviewer is unaware of the basis for that convention.
47	28-31	It is common practice to assign a default amount when the respondent cannot. This could also be imputed from the amounts that were reported. It is also common practice to assign "species, not specified." For species not specified, it would still be possible to determine if the fish was "caught by you or someone you know" and if so, if it came from Idaho waters or not.
48	7	This is incorrect. If they didn't know, they were assumed to be non-anglers.
48	19-20	Please confirm that this was true in most instances. It seems like the distribution of almost everything in this report would be skewed. Please be more specific.
51	Column headings	What does "n_w" mean? Please explain in a footnote.
51	Text box	How can it be that fish consumers are equally like to be male or female while fish non-consumers are more likely to be women than men? Isn't everyone either a fish consumer or not?
54	Text box	Please confirm that consuming 2,500 fish meals and snacks in 1 year was judged to be plausible. That's an average of 8 per day.
54	Text box	The median is better than what? The 99th percentile is not a measure of central tendency. The median is generally considered a better measure of central tendency than the mean for skewed data, as is the case here, but the mean is not mentioned in this paragraph.
63		"Snack" is not a place. Consumption events are either meals or snacks. This analysis is inappropriate and should be omitted.
69		If someone didn't eat any fish at the specified place, their portion size was counted as 0; but if they did eat at the specified place but didn't report their portion size, they were omitted from the analysis (hence the varying n's), right? This does not make sense. Please omit this analysis.
70	Text box, para. 2	This is an incorrect interpretation of the table. Why was the 65th pctl calculated?
80	4-6	The evidence provided does not support this statement.
81	6-9	Please reconsider re-weighting the sample to fix the income distribution; that will help with the proportion of anglers, too. Is it possible to re-weight the sample to make it more representative of Idaho without worrying so about the health districts? IDEQ does not need estimates at the health district level.
99	10-25	Elsewhere in this report it says that all of the 7 days that had fish were asked about, but this set of questions only asks about the most recent day that had fish. If additional days were asked about, please add the

Specific Observations and Comments on "Idaho Fish Consumption Survey"		
Page Number	**Line Number**	**Comment or Question**
		question that asked about another day with fish and the skip pattern(s).
117	Text box, para. 1	This does not seem to be true because the questionnaire does not indicate that questions about fish consumption were asked for more than "yesterday" and one additional day. Other results suggest that this was the case as well (Figures 13 and 14 and Tables 39 and 40). Please check the data file and see any if respondent has more 2 days of fish intake reported.

Review By:
Alanna J. Moshfegh, MS, RD

Peer Review Comments on "Idaho Fish Consumption Survey"

Alanna J. Moshfegh, MS, RD
Food Surveys Research Group, Agricultural Research Service, US Department of Agriculture
January 18, 2016

I. GENERAL IMPRESSIONS

This report provides a comprehensive and detailed description of design, execution, and findings of a research study on fish consumption of adults 18 years of age and older that reside in Idaho. Anglers were specifically over-sampled (33% of sample) in the study design. The information presented is based on a representative fish consumption survey of this sample. The information presented is extensive with thoroughness in the details of the sample design, data collection methodologies design, testing, and execution, and analyses of that data. Each step and consideration in the sample design and data collection execution are described giving the reader a thorough understanding and appreciation that nothing was left to chance and the study team did not fail to address issues and limitations related to this type of study. The information presented is quite clear in presentation and conclusions reached are supported by the data.

II. RESPONSE TO CHARGE QUESTIONS

1. Please comment the clarity and organization of the report. Does it present information, including tables and figures, in a clear and usable format? If not, please provide suggestions for improving the clarity of the document, which is intended to be useful to state regulators, the scientific community, and other stakeholders, including the general public.

The report is quite clearly written with easy to understand, clear graphics that help the reader to quickly realize the major results. Each section is labeled with detailed sub-sections for ease in locating information and understanding details. There are some minor suggestions I offer for improving clarity or discussion which I have detailed under "Specific Observations."

2. Please comment on the appropriateness of the survey design and implementation methodologies, including the food frequency questions and dietary recall questions, for the intended purpose?

The most current methodology for food frequency and 24-hour dietary recall were used and aptly adapted for collection of fish consumption only in this study. It is commendable that in addition to a 24-hour dietary recall, the contractor added a 7-day food record to ascertain intake of fish beyond just one day but up to 8 days. Details about pilot testing, interviewer training, and monitoring of data collection on 10% of the sample reflect a strong survey operation.

3. Please comment on the analysis and findings for differences in the portion size estimates between food frequency and dietary recall questions.

The differences found between food frequency and dietary recall on fish consumption is not unlike other studies. Food frequency generally overestimates number of times an item is consumed and is also less reliable on portion size estimation of compared to 24-hour recalls.

Recalls of very limited number of days may underestimate times an item is consumed, especially for those foods not typically consumed daily, such as fish. Thus, the follow-up recontact survey provides critical data to allow for usual intake analysis of fish consumption.

4. Are the results of the fish consumption survey scientifically sound and "valid" for use, in conjunction with the NCI method, for derivation of water quality criteria to be protective of human health for the general population and recreational anglers?

Yes, the results are sound and valid for use.

5. Do you have any other suggestions for improving the scientific quality or utility of the document?

Specific comments provided below.

III. SPECIFIC OBSERVATIONS

Specific Observations on Idaho Fish Consumption Survey		
Page	**Paragraph**	**Comment or Question**
3-7		Detailed Table of Contents excellent for finding information in report.
9	Line 11&29	Why not report same amt of fish by food frequency collection in these two places?
8	Line 4	Why not give the same type of figure found in this study somewhere in the summary? Is that figure the 31.23 gm reported in page 9, line 16? I am assuming the basis of that number is not the same as the 17.5 gm in line 4.
11	Line 7	First sentence does not make sense. Something is missing.
30		Excellent to have response rates by season for this research.
31		Response rates were low, but expected in a telephone survey. I did not see, but it might be helpful to include length of time of the interviews—does give appreciation to those who completed the survey.
43&		
49		Figure 2 results were nicely consistent with the 2012 Behavioral Risk Factor Survey.
54		Figure 5. Suggest giving average for fish consumers in the discussion.
56		Figure 6. Any discussion as to why number of meals with fish so much higher in October for anglers—popular/good fishing time???
57		Figure 7. Designation of 51-300 for heavy seems too broad a range, could ranges/further discussion be provided?

Review By:
Janet A. Tooze, Ph.D., M.P.H.

Peer Review Comments on "Idaho Fish Consumption Survey"

Janet A. Tooze, Ph.D., M.P.H.
Wake Forest School of Medicine
November 16, 2015

I. GENERAL IMPRESSIONS

In general, the methodology used to develop the surveys seemed sound, and the implementation seemed reasonable, with the possible exception of the low response rate. While these studies are generally prone to error, it appears that several steps were taken to minimize error. It appears that the survey weights were appropriately calculated from the survey. It also appears that the data were properly processed. I would have liked to have seen more discussion regarding whether these estimates represent usual dietary intake, particularly from the recalls. It is not clear if anyone with a nutrition background was involved in this process. Ideally a registered dietitian would have conducted the training and been involved in the design and interpretation of the analysis. It is concerning that the probability of reporting consumption on recalls decreases over time, and the impact of this was not addressed in a sensitivity analysis in the report. Statistical methods in general are not well described. There is no discussion of why normality was assumed with skewed data. It is not always clear if non-consumers are included or excluded in some statements. Some confidence intervals do not appear to be reasonable. In general, I would have liked more synthesis of all the information in this report, and more sensitivity analysis exploring differences between methods and assumptions made. For example, did the type of food (mixed dish or not), impact the portion size estimation? What happens if not all days of recall were used? I am also not sure of the reason for not trying to integrate the two reports (this one and the IMS report), but I think this really would help to improve clarity of the analysis of this survey. In addition, it may be useful to have IMS analyze some of the FFQ data as well as the recalls.

II. RESPONSE TO CHARGE QUESTIONS

1. Please comment the clarity and organization of the report. Does it present information, including tables and figures, in a clear and usable format? If not, please provide suggestions for improving the clarity of the document, which is intended to be useful to state regulators, the scientific community, and other stakeholders, including the general public.

The sections of the report are well organized, and results are presented in a logical format. However, the details as to which methods were used to test statistical significance are lacking from the tables (and even in much of the text). The tables should stand alone, and should mention the statistical methods used to determine statistical significance, and should include sample size (some do, but not all). Also the table titles should clearly indicate the source and the time period. Some tables like Table 17 have subscripts with no explanation of what they mean.

2. Please comment on the appropriateness of the survey design and implementation methodologies, including the food frequency questions and dietary recall questions, for the intended purpose?

The methodologies used were appropriate for the intended purpose of measuring dietary intake. Research has shown that food frequency questionnaires (FFQs) are subject to substantial measurement error, including both random and systematic error (bias). This is because they are quite cognitively challenging, requiring people to do averaging in their heads over a fairly long time period. One strength of this study was the cognitive testing and pilot testing of the questionnaires. These procedures would be expected to minimize errors with clarity of the questions and consistency of administration. All surveys were administered by phone, which is generally considered to yield similar results to in-person administration, and the pilot testing done by NWRG supported this as well, so that seems to be a valid way to administer the surveys. The cognitive testing supported asking participants to recall fish consumption by location (home, restaurant, other) and by meal, which hopefully minimized errors in recall. Because the FFQ was only focused on fish consumption, it also queried a large number of mixed dishes that contained fish, which should lead to fewer assumptions that need to be made regarding recipe content, which is also a strength. Participants were not asked about seasonal fish consumption or fish consumption at special events or to specify amounts for different types of fish, which might have improved estimation of intakes.

An 8-day recall was used to assess fish intake as well. This method is not commonly used, but would be expected to have errors similar to a 24-hour recall, which is more commonly used to assess short term intake. In general, short-term recalls are also prone to both random and systematic error like FFQs, but recalls tend to have less systematic error than FFQs.

Furthermore, statistical methods such as the NCI method may be used to obtain estimates of the distribution of usual intake that are adjusted for random error. Figure 13 is of concern, that recall of fish intake went down as time from the interview increased. This suggests a systematic error in the 8-d recall that would be expected to lead to an underestimate of true fish consumption, as one would expect the 'yesterday' recall to be the most accurate. It would be helpful to do some analyses where intake is normalized to the first day and sensitivity analyses of this issue.

3. Please comment on the analysis and findings for differences in the portion size estimates between food frequency and dietary recall questions.

I think that the reasons for the differences between the two methods include:

(1) The averaging component that is required for the FFQ. Participants must average across the past 12 months and across mixed dishes and non-mixed fish consumption. This is quite challenging, and might lead to smaller portion sizes to balance mixed dishes from non-mixed.

(2) The recall asks for consumption of a specific meal and type of fish, which does not require averaging and would be expected to be more accurate.

(3) The FFQ asked about "number of events" and the portion size of each event (meal), whereas the recalls asked about the "number of helpings" at each event and the portion size of each helping. This might result in larger portion size estimation as well.

(4) One point I wanted clarification on is whether it is true that all the portion size questions where they were asked to provide a range coded both "between 13 and 16 oz" and "more than 16 oz or more than one lb" as "06" (see for example FFQ5A_1 and FCR24_7B_1, but it seems to be true for all frequency questions)? Of course if both were given the same code in the database this could lead to inaccuracies for the FFQ and recall.

(5) The means in the report are for skewed data, and therefore the mean is not the best measure of central tendency of the distribution.

(6) There is no measurement error adjustment. It would be interesting to compare to the consumption day amounts from the NCI method with these estimates, but this does not appear to have been done.

4. Are the results of the fish consumption survey scientifically sound and "valid" for use, in conjunction with the NCI method, for derivation of water quality criteria to be protective of human health for the general population and recreational anglers?

There are limitations of these data. First, they rely on self-report and methods that are known to be subject to error. Second, the response rate was low, with only 25% of potential participants participating. The participation rate of 40% was better, but there is still potential for non-response bias. There is no adjustment for measurement error in this report. However, techniques were implemented to reduce potential bias in the surveys. Overall, I would conclude that the methodology used to implement this survey was scientifically sound, and the results are informative. I would rely more on the NCI method results and the recall results (particularly the first day), rather than the FFQ results to provide "valid" estimates.

5. Do you have any other suggestions for improving the scientific quality or utility of the document?

As I mentioned previously, I think it is important to describe the specific statistical tests that were used for each analysis. This is often not clear in the document. Sometimes statements are not properly cited in the text, or inadequate details are given to back up statements that are made. I have tried to specify many of these comments in the next section.

III. SPECIFIC OBSERVATIONS

Specific Observations on "Idaho Fish Consumption Survey"		
Page Number	**Line Number**	**Comment or Question**
8	3	HHC should be defined here.
9	4-19	It is not clear if these are rates for consumers only, or if they include the 12% non-consumers. This is also true in the results section.
10	3-5	This statement needs to be substantiated. How is it clear?
10	15-17	How was it assessed to be "significantly lower"? Did you take into account one asks about 1 year and one asks about 8 days? Did you compare the question on the FFQ that asked about the past 7 days to the recall? I'm afraid a statement like this is just going to confuse people – of course we would expect less consumption for a 1 week period compared to a 1 year period.
10	19-20	Why is this true? Is there a reason you would assume this?
11	28	The objective I believe is to assess usual dietary intake of fish. I think calling it a "daily rate" is not quite right – it is a usual rate expressed on a daily basis, but as a long run average.
14	10-12	Please state what these strengths and weaknesses are.
21	17	I do not think it is accurate to say that FFQs are representative of usual intake. This is not cited here.
21	18	I do not think FFQs are a preferable method for foods with day to day variability. It does not state to what they are preferred, and there is no citation for this.
21	19	I think this statement is debatable. Even highly literate subjects have trouble estimating year-long averages over multiple types of fish in their head. This is not cited.
23	5-7	Not being as dependent on memory as an FFQ is actually a strength of dietary recalls, particularly for 1 day. I would not cite it as a weakness, although there is some evidence of weakness for this multi-day recall used in this study.
48	20	Why were the assumptions made that the data were normally distributed when they appear to be so skewed?
49		Please include a 95% CI for the 88%.
51		What are the subscripts? What statistics were used?
54		This should specify a time period – annually.
56		What type of test was used?
71		These CI seem to be way too narrow. How were they computed?
80	4-6	How is it clear?
80	17	This seems low, as 50% is the standard for most survey studies. Please cite this. This is not a public opinion study.

ATTACHMENT A
Markup of "Idaho Fish Consumption Survey"
by
Patricia M. Guenther, Ph.D., RD