United States Environmental Protection Agency | Office of Water (4305) | 823B16001 June 2016

Guidance for Conducting Fish Consumption Surveys

DRAFT

DRAFT
Guidance for Conducting Fish Consumption Surveys

DRAFT

Office of Science and Technology
Standards and Health Protection Division
Office of Water
United States Environmental Protection Agency
Washington, DC

June 2016

GLOSSARY AND ABBREVIATIONS

Access Point Survey: A survey that is administered at locations where fishers gain entry to fishing or hunting areas. Examples include boat ramps, docks, and wildlife refuge check stations.

Accuracy: A measure of agreement, expressed numerically as a percentage, between a measured value and an accepted or true value.

Address-Based Sampling (ABS): A type of sample that uses addresses selected from the U.S. Postal Service Delivery Sequence File (DSF). In address-based samples, the address is the unit selected.

Aerial Survey: Flying over a fishing or hunting area to obtain an estimate of the total population participating in the activity during the period of time in which a creel survey or personal interviews are being conducted. This procedure is used to estimate the percentage of the population interviewed when other sampling strategies (e.g., probability sampling) cannot be used.

Area Probability Sample: A type of sample which uses geographic areas and usually includes clustered or a multi-stage sample selection.

Bias: Property of a statistical estimation process that consistently overestimates or underestimates a population parameter. The discrepancy between the expected value of an estimate and the population parameter being estimated.

CATI: Computer-Assisted Telephone Interviewing, a method of telephone interviewing in which a structured questionnaire is programmed into a computer. The interviewer enters the respondent's replies directly into the computer program.

Census: A complete enumeration of a population.

Commercially Caught Fish: Fish caught or harvested for commercial profit.

Confidence Interval: The range of values within which it is estimated a population parameter lies with a defined level of confidence based on sample data.

Confidence Level: The probability that a population parameter lies within a given range.

Consumer-Only FCR: Estimates of fish consumption rates that are for the subset of the study population (and the larger population they were selected to represent) that consume fish. They are calculated utilizing data from those study respondents that reported fish consumption, excluding those who did not. The time period for defining consumer-only can vary – 24-hr recall, 30-day consumption, any consumption in a year, etc., and this choice, along with the statistical methodology, affects the estimate of FCR. (See Per Capita.)

Creel Survey: A survey of anglers generally employed by fishery managers that collects data on the types of fish caught, the size of the fish, time spent fishing, and other data related to both the fishing trip and the catch.

Descriptive Statistics: The branch of statistics that involves summarizing, tabulating, organizing, and graphing data for the purpose of describing a sample of objects or individuals that have been measured or observed. Descriptive statistics typically include means, medians, standard deviations, percentages, and other statistics that illustrate magnitudes, diversity, and frequency.

Frequency Distribution: A tabular or graphical presentation of the number of times each value occurs in the data set.

Fish Consumption Rate (FCR): The quantity of fish consumed per unit of time, e.g., grams per day (g/day).

General Population: All individuals in a geographic area, without reference to any specific characteristic.

Heritage Rate: A heritage rate is the amount of fish consumed prior to non-indigenous or modern sources of contamination and interference with the natural lifecycle of fish. While it is often thought of as a historic rate, it can also be a current rate and an aspirational rate.

Inferential Statistics: The branch of statistics that involves making inferences about the value of one or more population parameters, on the basis of sample statistics. The most common applications of inferential statistical procedures are estimation and hypothesis testing.

Intercept Survey: A survey conducted in-person at a specified location of interest. The location may be specific to the population of interest. Selected participants are interviewed at the selected locations.

Iowa State University Method: A statistical program (Software for Intake Distribution Estimation -SIDE) to estimate the distributions of usual intake of nutrients, foods consumed almost daily, and other dietary components produced by researchers in the Department of Statistics at Iowa State University in 2001. (http://www.side.stat.iastate.edu/)

List Sample: A type of sample drawn from a list frame. List sample frames consist of households or persons who possess a characteristic unique to the list. Examples of list frames include persons with fishing licenses, tribal registries, etc.

Measurement Error: The difference between a measured value of a quantity and its true value.

Measures of Central Tendency: Descriptive statistics that identify the center or middle of a distribution. Common measures are the mean and median and, less commonly, the mode.

Measures of Dispersion: Descriptive statistics that identify the spread of values of numerical data. Common measures are the standard deviation, variance, and range.

Multimode Data Collection: A data collection protocol that utilizes more than one method for contacting and interviewing respondents. Multimode data collection methods are used to maximize survey response while controlling costs.

Multiple Source Method (MSM): A statistical program to calculate usual dietary intake from 24hr recall data and supporting data developed by the Department of Epidemiology of the German Institute of Human Nutrition Potsdam-Rehbrücke and is available for on-line analysis. (https://msm.dife.de/)

Multivariate Analysis: The analysis of data consisting of multiple variables and examination of associations among variables (e.g., regression and correlation analysis, analysis of variance and covariance).

NCI Method: A statistical method to estimate usual dietary intakes of foods and nutrients developed by the Nation Cancer Institute, National Institutes of Health, USA, and researchers at other institutions. This method can be used to estimate the distribution of usual intake for a population or subpopulation; assess the effects of individual covariates on consumption; and predict individual intake for use in a model to assess the relationship between diet and disease or other variable. (http://appliedresearch.cancer.gov/diet/usualintakes/method.html)

Oversampling: A process by which a specific group or target population is selected at a higher rate than other population members, in order to ensure estimates can be calculated for these specific groups.

Per Capita FCR: Estimates of fish consumption rates that are for the entire study population (and the larger population they were selected to represent). These estimates include those respondents who rarely, or perhaps never, consume fish. (See Consumer Only.)

Precision: A measure of uncertainty in an estimate, such as a standard error, coefficient of variation (CV), or the width or half-width of a confidence interval. A specified value of the precision is commonly used as a target in designing a survey, including specifying the survey's sample size. Following completion of the survey, each major estimate derived from the survey is usually presented along with a measure of its precision.

Random Digit Dialing (RDD): A method used to select samples for telephone surveys by random selection of telephone numbers within working exchanges. This method permits coverage of both listed and non-listed telephone numbers.

Random Error: A source of error that contributes variability (reduces precision) but does not influence the sample mean or median. It is a type of measurement error.

Response Rate: A rate indicating the proportion of cases completing specific components of a survey. In simplest terms, the number of completed cases divided by the total sample.

Recall Error: A response error resulting from a respondent's inaccurate recollection of particular events.

Rolling Cohort Method: A survey method that involves randomly placing survey participants into groups (cohorts), which are then sequentially surveyed over equally spaced intervals (e.g., 2 or more weeks). Each cohort is asked to provide recall data for a period of time equal to the interval spacing between cohort surveys. This method is typically used to provide coverage over an entire year while avoiding the problems associated with long recall periods.

Roving Creel Survey: A creel survey that is conducted by having the interviewer move through the survey area in a random or defined pattern to contact fishers.

Sample Size: The number of cases (usually households or persons) selected from a frame to be interviewed or asked to complete the survey. This is usually larger than the number of completed cases used in the analysis due to nonresponse and missing data.

Sampling Frame: A group of households or persons from which a survey sample is drawn.

Screening: The process of using a brief series of questions early in the survey process to identify whether sampled cases belong to the target population or possess a characteristic of interest.

Self-Caught Fish: Fish caught or harvested for personal consumption, or to share with family and friends. This is typically phrased as "fish caught by you or someone you know" in fish consumption surveys.

Serving Size: Amount of fish consumed at an individual meal or at a single point in time.

Statistical Program to Assess Dietary Exposure (SPADE): A statistical methodology that estimates the habitual intake distribution for daily and episodically consumed foods or dietary components based on information of intake, measured on a limited number of days (e.g. 24-hour recalls) and sometimes in combination with a questionnaire (e.g. food frequency questionnaires), produced by National Institute for Public Health and the Environment, Bilthoven, The Netherlands. (http://www.rivm.nl/en/Topics/S/SPADE/About_SPADE)

Sport Fishers: Fishers who fish for pleasure or competition.

Stratified Sample Design: A sampling design that separates population attributes into non-overlapping groups (strata) from which samples are to be selected. The establishment of strata occurs prior to sampling.

Subsistence: Relying on natural resources to provide for basic needs.

Suppression: The reduction in desired intake or consumption due to environmental or other factors beyond the control of a population (e.g. fears of chemical contamination in fish, fish populations of inadequate size to support consumption, loss of access to fisheries resources, loss of access to fishing equipment, changes in social structure affecting harvest). (See also "Heritage rates")

Systematic Error: A source of error in which measurements consistently depart from the true value in the same direction; affects the sample mean or median and can result in incorrect estimates and conclusions; another term for bias.

Usual Fish Consumption Rate (UFCR): Term used to denote an estimate that corresponds to a long-term average daily consumption rate.

Weights: Sampling weights are needed when sampled units are selected using unequal probability sampling and the use of these weights during analysis allows for the results to be generalized to the population form which the sample was drawn. Weights are used to adjust for nonresponse and to assign greater relative importance to some sampled elements than to others. Weights are calculated as the inverse of the probability of selection, then adjusted for other factors to improve accuracy.

DISCLAIMER

This guidance document provides advice on how to conduct surveys to estimate fish consumption rates. This guidance does not impose legally binding requirements on the U.S. Environmental Protection Agency (EPA), states, tribes, other regulatory authorities, or the regulated community, and may not apply to a particular situation based upon the circumstances. EPA, state, tribal, and other decision makers retain the discretion to adopt approaches on a case-by-case basis that may differ from those in the guidance where appropriate. EPA may update this guidance in the future as new or additional information becomes available.

The Office of Science and Technology, Office of Water, U.S. Environmental Protection Agency has approved this guidance for publication. Mention of trade names, products, or services does not convey and should not be interpreted as conveying official EPA approval, endorsement, or recommendation for use.

The suggested citation for this document is:

U.S. EPA. (2016). *Guidance for Conducting Fish Consumption Surveys*. EPA 823-B1-6001.Washington, DC: U.S. Environmental Protection Agency, Office of Water.

SUMMARY OF INTENT

The U.S. Environmental Protection Agency (EPA) has prepared this guidance document for fish consumption surveys in an effort to assist tribes, states, local governments, and others in designing and conducting statistically valid fish consumption surveys with valid analytic results. EPA recognizes that studies of finfish and shellfish (hereafter, referred to as "fish") consumption patterns are essential in developing water quality standards and assessing the health risks posed by contaminants in fish. These studies also play an integral role in developing advisories and bans to ameliorate these risks, and to measure the level of fishing activity associated with a particular body of water as well as harvest patterns. Studies are often focused on geographical or cultural populations potentially at high risk. Surveys of selected populations are used to estimate how much fish tissue is consumed and the frequency at which it is consumed. Data on exposure and determination of the distribution of average daily intake are necessary to assess risks posed to consumers of fish and shellfish (U.S. EPA, 2000b). Consumption patterns, including the types and amounts of fish consumed, parts of fish consumed, seasonality of fish consumption, frequencies of meals eaten from these organisms, and the preparation methods used, can also vary greatly within populations because of differences in age, gender, cultural practices and/or socioeconomic status. Information obtained from fish surveys can be used to determine whether the amounts of fish being eaten are safe in relation to possible chemical contamination, to estimate risks to persons who could consume fish that might contain bioaccumulative and potentially dangerous levels of toxicants, to develop water quality standards to protect human health, and to assess the effectiveness of fish advisories. The design of surveys is intimately linked with survey objectives. Care must be taken in utilizing the results of surveys designed for one purpose for other objectives.

The purpose of this document is to provide guidance for the design, conduct, and analysis of surveys focused on characterizing contemporary ingestion of finfish and shellfish. Building on the previous EPA guidance document (U.S. EPA, 1998), the discussion of survey methodologies has been updated from the previous document to reflect more recent developments in the area of survey research, including use of cellular telephones, the web, mobile devices, and use of multi-mode data collection designs. To supplement and provide context to the described approaches, this guidance document also covers a broad overview of the numerous and complex issues surrounding the development of a study approach, identification of survey objectives, sampling options, mode selection, questionnaire development, and operational and analytic considerations. New sections on the topics of consumption suppression and the role of heritage rates, especially among tribal populations, have been added. Also, in recognition of the fact that resources for fish consumption surveys can typically be limited, this document addresses survey design options within the context of budgetary resources to help the researcher make choices that best fit the situation. This document does not provide direct guidance focused on collection and analysis of fish tissue for contaminants or conducting surveys to assess understanding of and compliance with regulations or voluntary programs even if the surveys are relevant to water bodies with consumption advisories.

This guidance document is written for users with at least a basic familiarity of statistical and survey research techniques. References are provided throughout the document to help readers understand the important elements of topics that may not be their primary area of expertise. However, it should be noted that more complex survey designs and analyses required to address more challenging research questions or advanced analytical goals may require the expertise of a survey and/or

statistical professional. Consultants in both areas are widely available and may indeed be found in other government departments, academia, or in survey research organizations.

There are many considerations when designing and implementing a survey. In particular, planning for monitoring response rates and handling survey nonresponse are important issues. Nonresponse can have a varying degree of biasing effect on survey results (as discussed later in this document). Topics that are not included in this guidance document include: detailed methods for achieving higher response rates; various approaches to the use of survey incentives; procedures for hiring and training qualified interviewers; monitoring study progress; and other procedures that affect survey quality. A survey operations specialist could provide assistance to ensure that all appropriate steps are being taken to design and implement a high quality data collection effort that will satisfy the stated goals and objectives of the planned research.

TABLE OF CONTENTS

CONTENTS (continued)

CONTENTS (continued)

CONTENTS (continued)

CONTENTS (continued)

Tables

Figures

1 BACKGROUND

1.1 Purpose and Objective

Concern over potential human health risks associated with chemically contaminated fish (Ahmed et al., 1993; National Research Council, 2000; U.S. EPA. 2009; Stahl et al., 2013; Stahl et al., 2014; Wathen et al., 2015; Thompson & Boekelheide, 2013) has led many tribal, regional, state, and local governments to study fish consumption patterns among their citizens. Surveys are useful for deriving information about present, recent past, and near future fish consumption rates. The suitability of using surveys to gather information about past consumption rates depends on how long the baseline fish consumption rates have been altered or suppressed. For example, a recent contamination event that alters fish harvest and consumption patterns is amenable to before-and-after event comparisons. However, if baseline fish consumption patterns have been affected by prolonged contamination or suppression, a survey of contemporary people will likely not be able to define baseline (or heritage) consumption rates. See Chapter 6 for a discussion of non-survey approaches for assessing heritage rates and estimating suppression of fish consumption.

The processes and procedures by which government agencies develop consumption surveys vary widely. Additionally, the survey results are used for a variety of purposes, such as a basis for setting water quality standards, evaluating health risks posed by contaminants in fish, evaluating harvest rates or patterns, or assessing the effectiveness of consumption advisories or bans. In an effort to assist tribes, states, local governments, and others in designing and conducting statistically valid fish consumption surveys with valid analytic results, the U.S. Environmental Protection Agency (EPA) has prepared this guidance document for conducting contemporary fish consumption surveys.

The purpose of this document is to provide guidance for the design, conduct, and analysis of surveys focused on characterizing contemporary ingestion of finfish and shellfish. Hereafter, the term "fish" is used to represent both finfish and shellfish, unless otherwise stated. These surveys may be intended to characterize the fish consumption rates of the full population of the locality or of targeted groups such as recreational or subsistence fishers, high-consuming individuals, pregnant or lactating women, women of childbearing age, or disadvantaged economic groups. This document does not provide direct guidance focused on collection and analysis of fish tissue for contaminants or conducting surveys to assess understanding of and compliance with regulations or voluntary programs even if the surveys are relevant to water bodies with consumption advisories.

Fish consumption rate data are essential in developing water quality standards, and they also play an integral role in developing consumption advisories and bans and in legal proceedings. More broadly, data on contaminant exposure and determination of the distribution of long-term average intake are necessary to assess risks posed to consumers of fish (U.S. EPA, 2000b). Consumption patterns, including the types, parts, and amounts of fish, seasonality of fish consumption, frequencies of meals eaten, and the preparation methods used, can vary greatly within populations. These variations may be related to demographic characteristics, such as age, gender, race/ethnicity, socioeconomic status, and geographic area (Birch et al., 2014; Mahaffey et al., 2009; Connelly et al., 2012; Razzaghi & Tinker, 2014; Soon et al., 2014; U.S. EPA, 2014). Consumption patterns can also differ between populations because of differences in cultural practices (Ellis et al., 2014; Judd et al., 2004).

The survey methods presented in this document may be used to obtain information on the consumption of fish for purposes of estimating long-term average intake for various fish types of interest. Examples include all fish, locally caught fish, anadromous vs. resident fish, fish from a particular body of water, and fish from various habitat types (e.g., freshwater, estuarine, near coastal marine, marine). This information can then be used to determine whether the amounts of fish being eaten are safe in relation to assessed chemical contamination, estimate risks to persons who could consume fish that contain unhealthy levels of bio-accumulative toxic compounds, and set water quality standards.

1.2 Relationship of This Document to Other Guidance Documents

The EPA has developed a series of five documents designed to provide guidance to tribal, state, local, and regional environmental health officials who are responsible for issuing consumption advisories for non-commercially caught fish and shellfish. The documents are meant to provide guidance and do not constitute a regulatory requirement.

The first four documents are as follows (all are available at https://www.epa.gov/fish-tech/epa-guidance-developing-fish-advisories):

Guidance for Assessing Chemical Contaminant Data for Use in Fish Advisories,

Volume 1: Fish Sampling and Analysis (U.S. EPA, 2000a)
Volume 2: Risk Assessment and Fish Consumption Limits (U.S. EPA, 2000b)
Volume 3: Risk Management (U.S. EPA, 2000c)
Volume 4: Risk Communication (U.S. EPA, 1995)

In 1998, EPA developed a fifth document, "Guidance for Conducting Fish and Wildlife Consumption Surveys," concerning guidance on methods for obtaining consumption rate data for use in characterizing exposure in a population when estimating potential risks, determining whether a consumption advisory is warranted, and developing or modifying water quality standards (U.S. EPA, 1998).

Additionally, in 2000, EPA developed a document (U.S. EPA, 2000d), "Methodology for Deriving Ambient Water Quality Criteria for the Protection of Human Health (2000)." This document provides a methodology for developing human health ambient water quality criteria as required under the Clean Water Act. The methodology requires having an acceptable estimate of the rate of consumption of fish from fresh, estuarine, and near coastal marine waters for the population for whom the criteria are being developed to protect. In some cases the national default rate may be appropriate. In other situations, tribal, state-wide, or local rates are necessary. This document consists of three volumes and is available at https://www.epa.gov/wqc/human-health-water-quality-criteria.

This current guidance document updates the 1998 guidance document concerning methodologies for conducting fish consumption surveys by reflecting updated survey research tools and concepts,

and providing updates to the standard analytic methods for fish consumption surveys. It provides guidance that can be used to produce fish consumption rates for deriving ambient water quality criteria, among other uses.

1.3 Organization of This Document

There are five basic steps in the design and development of a survey for estimating fish consumption rates. Each of these steps will be addressed in this document.

1. Clear definition of survey objectives (including defining the population of interest)
2. Identification of specific information needs
3. Decisions about a survey approach and sample design
4. Development of an appropriate survey questionnaire
5. Consideration of implementation and operational issues

Chapter 2 presents a discussion of design considerations that the researcher must examine. These include articulating the overall objectives of the study, as well as understanding the population to be studied and the geographic area of interest. Outlining the information needs of the research is an important step, and this is also discussed in Chapter 2. Fish consumption behaviors, fish preparation methods, portion sizes, and seasonal variation in consumption patterns are all factors to be considered during the design phase. Issues surrounding suppression and heritage rates are also discussed as they pertain to comparing current levels of fish consumption with historical levels.

Chapter 3 presents a discussion of types of surveys, sample designs, and sample size requirements so the researcher will be aware of the numerous options that can be utilized, depending on specific survey requirements. Other issues surrounding sample design, including the target population, eligible participants, precision and accuracy requirements, stratification, and response rates, are also covered. This chapter also discusses challenges associated with some types of sampling frames. The chapter concludes with a series of tables indicating which approaches are appropriate/not appropriate/or may be appropriate—based on various survey objectives and types of populations. These options are presented within a budgetary framework designed to aid the researcher in determining which approaches will best fit their specific needs from both a scientific and a resource perspective.

Chapter 4 provides guidance on dietary collection methodologies using the approaches described in Chapter 3. Specifically, the reader is provided with a description of an existing instrument that can be used to collect data on fish consumption. Various factors that are relevant for instrument development, such as respondent recall periods, use of visual aids, portion sizes, etc., are also discussed. The use of qualitative methods and other pilot testing that can be used to finalize a survey instrument are described.

Chapter 5 presents survey implementation and operational considerations. During the development process, it is important to define eligible respondents, implement pre-contact procedures to maximize response rates, ensure the proper implementation of informed consent, and consider how other factors, such as language and literacy issues, will affect response rates and data quality.

Response rates are discussed and some methods for respondent retention (for surveys with multiple observation points) are provided.

Chapter 6 provides an overview of methodologies for measuring suppression and estimating heritage rates.

Chapter 7 discusses analytic considerations. Along with Appendix C, Chapter 7 provides a walk-through for using the National Cancer Institute Method (NCI Method) to analyze fish consumption data to obtain estimates of usual fish consumption rates.

Appendix A includes a discussion of additional considerations for instrument development to help the researcher develop questions that are methodologically sound, and provides example questions for use in existing surveys. Appendix B presents background on a variety of generally accepted dietary assessment methods. Appendix C provides details for the application of the NCI Method. Appendix D provides documentation of a publicly available, automated survey instrument along with a hard-copy version.

1.4 Research Design Issues for Fish Consumption Surveys

1.4.1 Development of Research Objectives

The design of a study needs to reflect the specific needs of the situation. For example, one study may be needed to determine an appropriate fish consumption rate (FCR) for updating a state's water quality criteria as required under Section 303 of the Clean Water Act (CWA). While EPA provides a national default FCR, because fish consumption varies by geographical location, racial/ethnic group, age, income, and possibly other factors, EPA suggests a "four preference hierarchy" for states and authorized tribes to follow in selecting a FCR to be used in the development of water quality criteria. The four preference hierarchy is as follows.

1. Use of local data
2. Use of data reflecting similar geography/population groups
3. Use of data from national surveys
4. Use of EPA's default intake rates

Local data may include data from a variety of contexts, including consumption by the general population state-wide, by a specific subpopulation within the state or region, consumption of fish taken from a specific water body or within a specific community, or a traditional baseline heritage rate. Depending on the data used, it may be appropriate to adjust the contemporary rate to account for suppression effects by documenting a heritage or unsuppressed rate with additional literature-based research (for tribes, for instance), or by evaluating recent past rates through a survey (see Chapter 6). Some government agencies may be interested in determining a rate to protect the most sensitive population(s) in a particular area, e.g., women of childbearing age. The national default intake is 22 g/day, which is the 90th percentile of freshwater and estuarine fish of the general adult U.S. population (U.S. EPA, 2014). In some circumstances a health-based rate may be desired such as

a rate of 32.5 g/day (8 oz./week) which is the amount of fish recommended in the FDA-EPA joint fish advice to obtain the health benefits from fish consumption (seehttps://www.epa.gov/fish-tech/epa-fda-advisory-mercury-fish-and-shellfish).

Each of the following factors will influence whether a survey is the most appropriate method, the survey design, including the sample selection to best represent the target population, the data collection mode (in-person, telephone, web, etc.), the survey instrument, and the analysis method.

While the purposes and uses of the data are established during survey design and questionnaire development, the manner in which the data are to be analyzed is not always pre-determined. To the extent possible, the details of analysis and interpretation methods should be defined early in the survey design process because they may have a significant bearing on the form and content of the questions to be asked. For example, using some analytical methodologies requires data gained through multiple contacts with participants and may require a minimum number of respondents with certain characteristics. Addressing these issues during initial planning and questionnaire design minimizes difficulties that may arise during data analysis and interpretation.

1.4.2 Budget Considerations

Surveys come in all shapes, sizes, complexities, and costs. Budget considerations play a significant role in the ultimate decisions about the design and conduct of any survey. Sampling frame development, quality control, data entry, data collection, and analysis can be costly and resource intensive. It is important to note, however, that low-cost resources available to conduct any given survey may include use of already acquired survey software, in-kind participation of government personnel (perhaps beyond the division sponsoring the survey), volunteer labor, and/or accessing existing government resources such as fish license registries. Alternatively, it may be possible to add a limited number of questions to an existing survey conducted by the state; for instance, the Behavioral Risk Factor Surveillance Survey (BRFSS) (http://www.cdc.gov/brfss/). Finally, the services of an outside consultant or consulting firm can be purchased. Costs of implementation may vary in different regions of the country.

For all these reasons, it is very difficult to provide estimates of the total cost of conducting any of the survey designs described in this document. However, within certain budget categories or ranges, this document provides guidance about a variety of methods that can be employed to ensure that quality fish consumption surveys can be performed given varying levels of available resources. Matching available resources with achievable survey objectives is perhaps the single most critical factor in the survey development/design process. Section 3.12 provides tables of survey approaches that are appropriate for various levels of funding.

1.4.3 Fish Consumption Rate (FCR) vs. Usual Fish Consumption Rate (UFCR)

An individual's FCR is intended to estimate the expected quantity of fish consumed per unit of time. For a population, there are a variety of FCR distributions depending upon which fish species are

being evaluated; for any particular distribution, some individuals consume more fish per unit of time and some consume less. Different time units can be used to express the rate (e.g., per day or per week). The FCR also depends on the window of time used in a questionnaire for querying fish consumption. The window might be one day or a year. The FCR is often estimated using statistical analysis. It may change over time. For example, it may be higher in the summer than the winter. Thus, the FCR depends on the time frame (e.g., summer, winter, annual) and the length of the time window used in asking about consumption.

There are three main types of FCR: 1) current, 2) baseline, and 3) heritage FCR. A current FCR is the amount of fish consumed per unit time that the target population is currently consuming. These rates are generally used in water quality criteria programs where the waters are not contaminated and fish consumption is not thought to be suppressed and for immediate interventions such as fish advisories. A baseline FCR is the amount of fish consumed prior to an event (i.e., a pre-release baseline) that has since decreased fish consumption, for example an environmental spill such as the Deepwater Horizon event. These rates are generally used in clean-up situations such as at Superfund sites where current rates are suppressed due to contamination of surface water and the site needs to be cleaned to so that consumption can safely return to the baseline rate. A heritage rate is the amount of fish consumed prior to non-indigenous or modern sources of contamination and interference with the natural lifecycle of fish. While it is often thought of as a historic rate, it can also be a current rate and an aspirational rate.

If the consumption of fish is infrequent or episodic, such as is observed in most of the U.S., the estimated FCR may be variable or imprecise. If fish consumption is of great cultural importance, recall estimates may be more precise. If a person eats fish for dinner every Friday and not at other times, the FCR is one fish meal per week and the estimated FCR is likely to be relatively constant. If a fish meal is consumed on average once every 7 days, but sometimes 3 days in a week and other times not for several weeks, the estimated FCR over a short time frame can be quite variable, even though the true FCR is constant and is the same as in the first example. As the time frame covered by the data gets longer, the estimated FCR becomes less variable. It is important to note that in some instances, even a time period of a year may not be a truly representative long-term average, due to variation caused by the year-to-year changes in the size of anadromous fish returns. Assuming the true long-term FCR is constant over time, if the time frame covered by the data is very long, the estimated FCR becomes a relatively precise estimate of the true long-term consumption. We can add the term "usual" to "fish consumption rate" (UCFR) to imply that the resulting estimates are those that correspond to long-term averages, rather than short-term estimates, and to avoid a distinction between the true rate and the estimated rate. For further discussion about usual intake, please see the National Cancer Institute's (NCI's) Applied Research website (http://appliedresearch.cancer.gov/diet/usualintakes/).

Assuming the long-term FCR is constant over time, and given that short-term dietary recall data are more accurate than long term recall data (Kipnis et al., 2003), methodologies can be designed to estimate the distribution of the true, long-term FCR even though the data are collected over a limited time frame.

In the mid-2000s, the NCI developed a statistical methodology to estimate usual intake of episodically consumed foods (Tooze et al., 2006; Tooze et al., 2010). This method, known as the "NCI Method," has been published, and statistical programs to implement it are available on NCI's website. There are also other commonly accepted analytical methods to estimate usual FCR. These

include use of a food frequency questionnaire (FFQ) and other statistical methodologies such as the Multiple Source Method (MSM) (Harttig et al., 2011), the Iowa State University Method (Nusser et al., 1996), and the Statistical Program for Age-Adjusted Dietary Assessment (SPADE) (Dekkers et al., 2014). These statistical methods, along with the NCI Method, have a general requirement that dietary data are collected through two or more contacts with participants utilizing a 24-hour recall to collect dietary data. The time between recalls should be long enough to assume that the consumption reported in the first is not related to (or correlated with) the consumption reported on the following 24-hour recall. The use of an FFQ to estimate usual fish intake is a lower cost option, as only one contact is needed with respondents. However, many details of dietary intake are not captured on FFQs, as each question in an FFQ typically represents several foods or includes all possible preparation methods. The portion sizes queried about in FFQs are generally categorized into three levels such as <2 oz., 2 to 6 oz., and >6 oz. Additionally, while both 24-hour recalls and FFQs are known to have measurement error, FFQs are more prone to systematic error (Kroke et al., 1999; Subar et al., 2003; Kipnis et al., 2003; Freedman et al., 2004) and for this reason they are not recommended by the NCI for estimating percentiles of usual intake (see http://dietassessmentprimer.cancer.gov/approach/table.html). For more information on measurement error see NCI's measurement error webinar series at http://epi.grants.cancer.gov/events/measurement-error/. Further discussion of these methods and a table depicting data needs and advantages and limitations can be found in Chapters 4 and 7.

An additional consideration is the distinction between consumer-only FCR and per capita FCR. On a theoretical level, a consumer-only FCR is the FCR for the population that consumes fish, excluding those who do not, while the per-capita FCR is the FCR for the entire population, both consumers and nonconsumers. In practice, it is difficult to determine who is a fish consumer and who is not. Is a fish consumer a respondent that reported fish consumption on the 24-hr recall? Is a fish consumer a respondent who reported they consume fish once a week, once a month, once a year? If a study is interested in the long-term average fish consumption (UFCR), then even those who rarely consume fish, say one time per year, should be considered fish consumers. The upper end of the resulting distribution of UFCR would represent the frequent consumers while the lower end of the distribution would represent the infrequent consumers. The NCI Method and other approximations assume all respondents included in the analysis to be fish consumers and each is assigned a non-zero probability of fish consumption based on the factors included in the model (e.g., gender, age, race/ethnicity, reported frequency of consumption). Some respondents have very low probabilities of consumption, and thus their resulting usual intakes are low. If a study is truly only interested in consumer-only FCR, then it must determine how it will define a consumer and include the appropriate question on the survey to exclude nonconsumers from the analysis.

1.4.4 Data Quality Objectives Process

The Data Quality Objectives (DQO) process provides a systematic procedure for defining the criteria that a data collection design should meet. It can assure that the type, quantity, and quality of data used in decision-making will be appropriate for the intended purpose. Following this process will require the survey team to focus on the purpose of the research and use of the findings before assuming that a survey will meet their needs as well as through the planning and designing of a survey. The DQO process has seven steps.

1. State the problem – describe the problem to be studied and review relevant literature and information
2. Identify the decision – identify what questions the study will attempt to resolve and what actions may result
3. Identify the inputs to the decision – identify the information that needs to be obtained and the measurements that need to be taken/data that needs to be collected to resolve the decision statement
4. Define the boundaries of the study – specify the time periods and spatial area to which decisions will apply and determine when and where the data should be collected
5. Develop a decision rule – define the statistical parameter of interest, specify the action level, and integrate the previous DQO outputs into a single statement that describes the logical basis for choosing among alternative actions
6. Specify tolerable limits on decision errors – define the decision maker's tolerable decision error rates based on a considerations of the consequences of making an incorrect decision
7. Optimize the design for obtaining data – evaluate information from the previous steps, generate alternative data collection designs, choose the most resource-effective design that meets all DQOs

For more information see Guidance for the Data Quality Objectives Process, U.S. EPA, 1994, available at https://archive.epa.gov/epawaste/hazard/web/pdf/epaqag4.pdf.

2 SURVEY OBJECTIVES AND INFORMATION NEEDS

2.1 Overview

As listed in Chapter 1, there are five basic steps in the design and development of a survey for estimating fish consumption rates. Each of these steps should be addressed within the researcher's budgetary and resource limitations.

This chapter addresses the first two of these steps – Clear definition of survey objectives (including defining the population of interest) and Identification of specific information needs.

2.2 Defining the Survey Objectives

A clear definition of the survey objectives is an integral step in the survey development process. Fish consumption estimates obtained through a survey for one set of objectives may not be appropriate to use for another purpose. For example, current FCR estimated for use in setting advisories may not be appropriate for setting water quality standards if fish consumption is not at baseline levels (i.e., unsuppressed rates). If used for this purpose, water quality standards would be set where unsuppressed consumption may result in unhealthy exposure. As each survey approach has unique advantages and disadvantages, the survey objectives should guide the conduct and design of the sample, survey questionnaire, and analysis plan to obtain specific results (e.g., estimated UFCR, estimated age distribution of consumers). The information collected must be targeted to address the objectives, and thus serve as a planning tool to ensure that the required information is collected in a manner consistent with planned analysis techniques. Through defining the objectives, the team may find that they require heritage rates, thus a survey is likely not the best approach (see Chapter 6).

Key considerations include:

- The planned use of the data (e.g., setting ambient water quality criteria, setting cleanup levels, assessing human health risk, assessing advisory effectiveness, risk management)
- The target population (e.g., general population of state, county, other location of concern, local fishers, women of childbearing age, tribal members, consumer-only vs. per capita, their history and lifestyle)
- The type of fish of interest (e.g., all fish consumed from all sources, estuarine and freshwater fish only, a specific fish species or group, locally caught or commercially available)
- The geographic area of concern (e.g., tribal lands, entire state, one particular lake or river)
- The condition of waters within the geographic area of concern (e.g., existing fish consumption advisories, watershed contamination, environmentally sensitive area)
- The timing of the survey period
- The targeted number of data collections from each participant

In state or tribal ambient water quality criteria programs, local fish consumption data are preferred over national default rates. Thus a state or tribal program may define the survey objectives as follows:

- Estimate the distributions of fish consumption for various fish types by habitat among the population of interest
- Estimate the proportion of fish consumption by fish habitat (i.e., marine, near coastal marine, estuarine, and freshwater and anadromous vs. resident) or by watershed
- Estimate the distribution of consumption of locally caught fish among high consumers
- Estimate the proportion of total fish intake that is locally caught vs. commercially available

Given these objectives, the questionnaire must collect information on the species consumed in order to obtain rates for estuarine, freshwater, near coastal marine, and marine fish separately from total fish and source of the fish. Additionally, the study needs to be representative of the population of interest and needs to ensure enough frequent consumers are represented in the sample. Weighting of the sample might be employed to adjust the representation of frequent consumers to reflect that of the population. It may not be possible for a state survey to accurately characterize the consumption levels of specific groups within the general population given the low fraction of survey results that are reflective of these groups. It will be particularly difficult to develop upper percentile estimates of fish consumption for such groups. To the degree possible, a state survey should try to include enough frequent consumers so that their contribution to the general population FCR is reflected.

Groups that are highly segregated from the general population or distrustful of governmental entities may not respond to a survey administered by a state agency. In these instances, it may be necessary for the state to utilize the results of additional surveys that are not managed by the state to characterize FCR for specific high fish consuming populations. The state can establish a relationship with the groups of interest, developing lines of communication and coordinating efforts regarding survey design and implementation to help to ensure that adequate and appropriate data from these groups are available to inform a state's regulatory efforts. Federal agencies with trust responsibility to tribes may assist in facilitating communication between states and tribes. See Section 2.2.2 for more information.

In some cases, consumption data are desired to evaluate effectiveness of advisories, i.e., the success of existing advisory messages recommending certain consumption behavior. Determination of actual consumption levels can serve to improve the accuracy of the advisory. For example, an advisory program assessing its message to limit consumption of largemouth bass caught at a specific lake to two times a month and advising the fat be trimmed from the fish and internal organs not be consumed, may define the survey objectives as follows.

1. Identify people who consume fish from the specific lake
2. Estimate the frequency of their consumption of largemouth bass caught at the lake
3. Estimate the distribution of long-term average intake of largemouth bass caught at the lake
4. Determine the parts of the fish that are consumed

Given these objectives, it is clear that a state-wide general population survey would not provide the data required and information about preparation of the fish would need to be included in the survey

instrument. Both the target population and the water body are relatively small, so the survey would best be conducted at access sites rather than through mail or telephone. Unless addresses and/or telephone numbers of the target populations are known, such approaches would be unlikely to capture enough respondents in the specific population for a statistically valid estimate.

2.2.1 Defining the Population of Interest

Identification of the population of interest should be articulated by the study objectives. Defining the population is of key importance in developing a sample that directly represents the population or which can be made to represent the population using strata and pre-considered weighting factors. Surveys can be designed to target groups that might be at greater risk of exposure to contaminants in fish due to higher consumption rates, such as subsistence fishers. Obtaining accurate exposure information (e.g., fish consumption data), is a critical aspect of this responsibility. Surveys can be designed to focus on populations that may be more susceptible to the health effects of contaminant exposure such as children, pregnant women, people with pre-existing health problems, the elderly, etc. If the target population is relatively small and it is desirable and feasible to survey the entire population (i.e., take a census), then the results obtained will be observations of the population parameters. More likely, however, is a situation in which a subset of the target population needs to be sampled; the results obtained are sample statistics which, if obtained correctly, are expected to be good approximations of the population parameters. Census estimates have less risk of error than sample estimates because they are subject only to the reliability, validity, and measurement error involved in the survey response (see discussion of Accuracy in Section 3.7). Sample estimates are subject to the same types of error risk as a census, but also to sample selection bias and sampling error risks.

States or others planning fish consumption surveys that could result in environmental regulatory actions, such as establishing water quality standards, for geographic areas that include tribal lands or populations should consult with EPA concerning how the survey and regulatory actions relate to any tribal treaty rights of the tribes affected. Under the Constitution, treaties have the same legal force as federal statutes. Many treaties reserve specific rights that relate to the environment by addressing a tribe's right to maintain their traditional ways of life, including rights to hunt, fish, and gather. Significantly, such reserved rights exist both on reservations and off the reservations in ceded territories under state jurisdiction. However, not all tribes have treaties. In these cases a state or federal consultation policy may assist the survey team in identifying appropriate tribes or aboriginal groups and ascertaining their legal status.

A major consideration is the size of the population of interest (e.g., number of pregnant women, or high fish consumers) and the prevalence of the behavior(s) of interest (e.g., consumption of locally caught fish) within the larger population. If the subgroup of interest or the behavior to be studied is relatively rare – or hard to find – then some sample designs should be eliminated from consideration. For example, a random digit dialing survey, address-based sample, or area probability sample may require prohibitively large screening efforts to identify a sufficient number of respondents meeting the criteria for the study. It is also important to learn as much as possible about the characteristics of the population of interest. The survey design, including sampling methods and mode(s) of administration, must carefully consider who the potential respondents will be. Questions such as the following should be addressed during the design phase.

- What is the population (or sub-population) of interest?
- Where do members of the population live, or where can they be found?
- What are the literacy and web penetration rates among the population of interest and what are the common methods for finding and contacting the members of the population?
- How many of them have telephone numbers (cell and/or landline) and/or fixed physical or mailing addresses?
- How many of them use land lines versus cell phones?
- Do the survey objectives require identifying populations of relatively low prevalence in the general population (e.g., subsistence anglers or pregnant women)?
- What response rates are needed in order to provide a large enough final expected sample size for certain subgroups to answer the study's main questions with sufficient precision?

If key characteristics of the population to be surveyed are known, survey designers should take these into account and rule out any sample designs or modes that are known to be unsuitable or inapplicable for a substantial proportion of the target population. For example, one should not plan to do a telephone survey on a group that does not include a high percentage of individuals with available telephone numbers. Similarly, a self-administered mail survey would not be suitable for a group with known low literacy rates or lack of fixed physical or mailing addresses.

2.2.2 Community/Tribal Input

For all surveys, procedures should be implemented for identifying, contacting, and coordinating with the appropriate community/tribal leaders early in the study design process. For state surveys that are also incorporating tribal populations or other culturally diverse sub-populations, this is an important step that can engage the tribe or local community and obtain their input into the research process. Tribal officials, local officials, and community partners can help frame the purpose of the study, and formulate objectives that will result in meaningful information for use by tribal members and other local residents or fishers. The tribal members can also be helpful in providing local, or traditional, names for fish, parts of fish consumed, unique fish preparations or dishes, fishing locations, feasts or communal meals where fish are shared, and perspectives on current vs. historic consumption. Local context is especially important among tribal populations. If the survey is not being led by the tribe, care should be exercised to ensure that tribal input and involvement is solicited at each step of the development process.

Community engagement efforts can serve several purposes. First and foremost, communicating with community leaders can help obtain cooperation from members of the community and ensure that the research meets the needs of the community. Engaging with community leaders can also help identify cultural norms that may be important for cooperation or collecting accurate data (Donatuto & Harper, 2008). For study designs that target specific populations such as particular ethnic groups (e.g., Hmong Americans) or American Indian and Alaskan Native tribes, where there may be inadequate sampling frames, input from community leaders can help to identify strategies for hard-to-reach community members.

As noted in Section 2.2, it may be preferable to both the tribe and the state, to have tribes (or other groups) conduct their own separate surveys to better accommodate collection of the detailed information that is associated with a smaller group for which fish consumption is an important cultural and/or spiritual activity. Characterization of the cultural, spiritual, and/or historic aspects of fish consumption, as well as the relationship between public health and fish consumption, are survey objectives that may be of greater importance to tribes than states. Additionally, survey modalities suited to the general population (e.g. mail, phone, internet) may not be suitable for all populations. Regulatory perspectives of tribal populations may differ from those of government entities and intended uses of the data should be clarified up front to avoid any misperceptions.

The survey team may be interested in using community-based participatory research (CBPR).CBPR is a partnership approach to research that involves community members, practitioners, and academic researchers in all aspects of the research process. It enables all partners to contribute their expertise and share responsibility and ownership (Israel et al., 1998). See Israel et al., 2012, Israel et al.,1998, and Minker and Wallerstein, 2003, for more information.

2.2.3 Geographic Area(s) of Interest

If fish consumption rates are desired for a specific geographic area of interest, the area in which the survey is to be conducted should be carefully defined. It will be a factor in how the sampling frame is developed. Is there an interest in surveying only persons living within a certain area regardless of where they catch their fish, or should the sample include persons who harvest and eat fish caught in a certain area? Defining the geographic area can be done on a watershed basis, a certain radius around a Superfund site, by travel distance to fishing locations, or by jurisdictional boundary. It may be important to review the history of the watershed, looking for past or current contamination, dams, spawning areas, environmentally sensitive areas, etc.

2.3 Information Needs

In addition to the overall purpose and objectives of a consumption study, the need for information about specific aspects of consumption or characteristics of fish consumers should be considered. The extent to which these factors are important and the type of information needed to meet the objectives of the study will influence which survey approach is selected.

2.3.1 Physical and Demographic Characteristics of the Population

Collecting physical and sociodemographic characteristics of fish consumers is important for a variety of reasons. These data allow for presentation of results by population characteristics and they allow for analysis of nonresponse bias. Characterizing fish consumption rates by factors known to be related to fish consumption in the presentation of results can increase the usefulness of the final UFCR estimates. For example, fish consumption is related to age, racial/ethnic background, gender, and income (Birch et al., 2014; Burger et al., 1999; Connelly et al., 2012; Holloman & Newman,

2010; Mahaffey et al., 2009; Razzaghi & Tinker, 2014; Soon et al., 2014; U.S. EPA, 2014). Holloman and Newman, 2010, found that low-income African American women in Newport News, Virginia, consume seafood at a subsistence fisher rate, even though they were not subsistence fishers. Body weight may be important data to collect in order to estimate FCR normalized to body weight (g fish consumed/kg body weight). Physical and sociodemographic characteristics of fishers that may be important (depending on the study objectives) include: race, ethnicity, gender, date of birth, height, weight, pregnancy/lactation status for women, physical disabilities or medical conditions, number of household members, occupation/employment status, income and education level, language spoken at home, and urban/rural residence. The extent to which potential predictors of consumption can be incorporated into the survey and used in the presentation of results depends on the number of questions that can be accommodated in the chosen survey design and the budget.

2.3.2 Hard-to-Survey Populations

Populations may be hard to survey because, among other reasons, members of the population may be: (1) hard to define or identify and thus to sample; (2) difficult to find or contact; and/or (3) unlikely to cooperate with surveys. Survey errors that may occur with hard-to-survey populations may be the result of sampling challenges (which may require special sample frames), noncoverage or undercoverage (the sampling frame does not include all members of the population), nonresponse (failure to obtain complete data from all selected individuals), and measurement errors (the difference between the "true" value and the reported or measured value) (Smith, T., in Tourangeau et al., 2014).

Populations or sub-populations are often hard to identify when there are no reliable lists of population members and/or when there are few individuals among the larger population who meet the criteria of interest for the survey. They may be hard to find if they have high rates of mobility or have barriers that impede their accessibility. For example, barriers may include the need for interviewers to travel long distances to reach the groups of interest, the prevalence of gated communities within urban areas, and the use of caller ID (and refusal to answer unknown callers) by many members of the population. Examples of groups unlikely to cooperate include those with little or no interest in the subject matter, those who may be suspicious of the survey organization or whomever it represents, or those with low investment in the community. Hard-to-survey populations are often defined by multiple factors such as these (see Tourangeau et al., 2014). Throughout this document, we have emphasized the importance of being as knowledgeable as possible about the population of interest. This will enable the survey design to incorporate techniques to minimize the impact that hard-to-survey groups may have on survey results.

For fish consumption surveys, special consideration of sampling challenges should be given when planning surveys of remote, possibly isolated populations. For example, American Indian/Alaska Native populations have historically been undercounted (Tourangeau et al., 2014). Hard-to-reach areas have often turned into unsampled areas due to their remoteness. Some locations in Alaska, for example, are only serviced by float plane (Smith, T., in Tourangeau et al., 2014). Telephone surveys may not be an option due to lack of coverage in these rural or frontier areas. For on-the-ground interviewers, difficulty in locating households may be encountered due to streets with no names or house numbers. This makes it difficult to verify sampled housing units. In addition, residents of these areas may have high rates of mobility, on a regular or semi-regular basis, or tenuous

attachment to a household. Not only is this problematic for identifying household members, but also for conducting followup interviews to obtain fish consumption data on multiple occasions, as required by some survey methodologies. For household surveys that require an enumeration of all household members, rules for identifying household members must be clearly defined and conveyed to field interviewers so they will know who is to be counted as a household member, whether or not the individual is physically present at the time of screening.

2.3.3 Fishing Activities and Behavior

Seasonal variability in the amount and/or type of fish consumed by the population of interest should be considered during the design process. This variability may be due to the availability of fish, fish spawning patterns, weather, or local customs or traditions associated with different holidays or times of the year. Seasonal variation may be a factor not only in the amount of fish consumed, but also in the type(s) of fish eaten, and/or in how the fish is prepared. For example, during the summer months, fish may be eaten fresh most of the time, but during the winter months it may only be eaten in dried form.

The direct approach for capturing seasonal variations and accounting for them in estimates of fish consumption is for data collection to span a full year. With direct estimation methods, a survey conducted during months of high availability of fish would be expected to overestimate annual consumption, and a survey conducted during months of low availability of fish would be expected to underestimate consumption. Ideally, equal proportions of the sample would be interviewed over fixed time periods until the entire sample had been interviewed. Unequal numbers of interviews per time period can be accommodated by weighting approaches. Another approach might be to conduct interviews during two periods, one of low consumption and one of high consumption. If it is not possible for the data collection period to span an entire year, indirect methods (which model the seasonal variation) may be used, but these approaches rely heavily on assumptions about the nature of seasonal variation (see discussion of usual fish consumption rates, or UFCR, in Section 1.2.3.) One approach that has been used to accommodate seasonal variation is to allow the respondent to designate two "seasons"—a high and a low consumption season—and report on typical consumption per species during each "season," along with the duration of the season (adding to 12 months). See, for example, the Suquamish Tribe, 2000.

Examples of other possible fishing-related activities and information that may be needed to meet survey objectives include the following:

- Location(s) of fishing activities (specific sites, type of water body)
- Distance(s) of fishing activities from principal residence
- Seasonal and temporal distribution of fishing activities (total number of days per season, which months of the year, overall or for each location)
- Fishing effort (hours/outing, hours/day, outings/month, days/month)
- Purpose for fishing (consumption, sport only: catch and release, etc.)
- Mode of fishing (e.g., nets, traps, hook and line, pier, shore, private boat, charter boat, scuba)
- Type of fish caught
- Number of fish caught per outing

- Size of fish caught
- How long involved in fishing activities and consuming self-caught fish (e.g., new to sport or years?)
- What was the fate of the caught fish (e.g., released, consumed by household, sold, given away?)

2.3.4 Preparation and Consumption Patterns

Fish preparation techniques may have a substantial impact on the levels of certain types of contaminant exposure (e.g., PCBs, dioxins, and DDT) and, therefore, can be used as an additional predictor of exposure levels (Forsberg et al., 2012; Salama et al., 1998; Skea, et al., 1979; Voiland et al., 1991; Zabik, et al., 1995; Zabik et al., 1996; Zabik & Zabik, 1999). Preparation methods include how the fish was cleaned or trimmed (e.g., was the fat trimmed away?), as well as how—or if—it was cooked (e.g., baked vs. fried). Response categories to questions about preparation methods should match local culinary practices and the types of preparation methods of interest for the study objectives. The change in weight of the fish due to cooking (moisture loss, fat gain, etc.) varies by cooking method. Thus, in order to convert as-consumed weights to raw weight, cooking methods need to be collected.

Examples of possible preparation and consumption pattern information that may be needed to meet survey objectives include the following:

- Source of fish consumed (e.g., locally caught, commercially obtained)
- Amounts of fish eaten per meal/day/week/month for each person in household (visual cues are helpful to improve the accuracy of portion size estimates)
- Geographic and seasonal variations in consumption
- Parts of fish consumed (may vary with the species)
- Parts of fish used for cooking but not ingested (e.g., boiling of bones or fish heads)
- How the fish were prepared for cooking/eating (e.g., skinned, fillet, steak, shucked)
- How the fish were cooked if they were not consumed raw (e.g., baked, fried, steamed, smoked)
- Fish consumed in mixed dishes such as fish stews
- Special cultural/ethnic practices in fish consumption and preservation (e.g., smoking, canning, drying)
- Consumption of fish obtained from restaurants, supermarkets, fish markets, roadside stands, from family or friends, or obtained in other ways such as tribal distribution programs
- Consumption of shared fish (e.g., shared in general or at communal meals, feasts, ceremonies, or other gatherings)
- Whether fish were frozen or preserved and eaten throughout the year, or eaten only when fresh

2.3.5 Suppression and Heritage Rate Issues

A "suppression effect" occurs when a fish consumption rate (FCR) for a given population, group, or tribe reflects a current level of consumption that is diminished from the level of consumption that the population, group, or tribe would actually consume in the absence of factors such as the following:

- Health-based suppression of consumption
  - Reduction or avoidance due to actual or perceived presence of chemical contaminants in the fish
  - Reduction or avoidance due to fish advisories
- Reduced fish populations due to environmental changes (e.g., changes in water conditions, alteration in fish habitat by dams and shoreline development)
- Reduced access to fish (due to fishing regulations such as creel limits or barriers to access such as land ownership and shoreline development)
- Changes in social structure; reallocation of time from traditional lifeways to other pursuits
- Imposition of laws or regulations reducing fish consumption

Suppressed fish consumption is of particular concern among some tribal populations since issues of physical, spiritual, or social health have been linked to fish consumption (O'Neill, 2013; Donatuto & Harper, 2008). Suppression may affect national fish consumption rates as well as rates among smaller populations, especially tribal groups. Among all populations, there may be the desire to be able to eat more fish than is available or safe to eat, for example the joint FDA-EPA fish advice advises people to consume 8 to 12 oz. of fish per week. EPA's "Human Health Ambient Water Quality Criteria and Fish Consumption Rates: Frequently Asked Questions" (U.S. EPA, 2013), states that when setting water quality criteria, "It is also important to avoid any suppression effect that may occur when a fish consumption rate for a given subpopulation reflects an artificially diminished level of consumption from an appropriate baseline level of consumption for that subpopulation because of a perception that fish are contaminated with pollutants." Environmental standards utilizing suppressed rates may contribute to a scenario in which future aquatic environments will support no better than suppressed rates.

A "heritage rate" is the amount of fish that was traditionally consumed prior to non-indigenous or modern sources of contamination and interference with the natural lifecycle of fish. A heritage rate of consumption is generally extrapolated or reconstructed from information available in anthropological or historical literature. In many cases, heritage rates may be the only practical way to estimate unsuppressed rates – that is, free from the biasing influence of suppression effects, and may be useful in establishing a baseline for legally protected fishing rights for fishing tribes. As discussed in Section 2.1.1, if the resulting FCR estimates from the survey will be used for regulatory action under the Clean Water Act, consultation with EPA is recommended.

Methodologies to assess if there is a suppression effect, to estimate the size of the effect, and to estimate heritage rates are presented in Chapter 6.

2.4 Environmental Justice

The U.S. EPA defines environmental justice as "the fair treatment and meaningful involvement of all people regardless of race, color, national origin, or income with respect to the development, implementation, and enforcement of environmental laws, regulations, and policies. EPA has this goal for all communities and persons across this Nation. It will be achieved when everyone enjoys the same degree of protection from environmental and health hazards and equal access to the decision-making process to have a healthy environment in which to live, learn, and work."
In 2002, the National Environmental Justice Advisory Council (NEJAC) published a document that discusses recommendations to EPA regarding environmental justice and fish consumption (NEJAC, 2002). The document and recommendations provided within address the question, "How should EPA improve the quality, quantity, and integrity of our Nation's aquatic ecosystems in order to protect the health and safety of people consuming or using fish, aquatic plants, and wildlife?" and should be considered carefully by the survey team. The six recommendations provided in the document are:

1. Require states, territories, and authorized tribes to **consider specific uses**, including the use of the waterbody or waterbody segment for subsistence fishing, when designating uses for a waterbody, and to set water quality criteria that support the specific designated use; provided that where human health criteria are established based upon consumption of toxic chemicals that bioaccumulate in fish, regulators should employ appropriate human fish consumption rates and bioaccumulation factors, including cultural practices (e.g., species, fish parts used, and manner of cooking and preparation) of tribes and other indigenous and environmental justice communities using the waterbody; provided further that EPA should encourage and provide financial and technical support for states, territories, and authorized tribes to control effectively all sources, including both point sources and nonpoint sources, to achieve the criteria
2. Work expeditiously to **prevent and reduce the generation and release of those contaminants** to the Nation's waters and air that pose the greatest risk of harm to human health and aquatic resources, including but not limited to persistent bioaccumulative toxics (PBTs) (e.g., mercury, dioxins, and polychlorinated biphenyls (PCBs)) and other toxic chemicals, and to clean up and restore aquatic ecosystems contaminated by pollutants
3. **Protect the health of populations with high exposure** to hazards from contaminated fish, aquatic organisms and plants, and wildlife, including communities of color, low income communities, tribes, and other indigenous peoples, by making full use of authorities under the federal environmental laws and accounting for the cultural, traditional, religious, historical, economic, and legal contexts in which these affected groups consume and use aquatic and terrestrial resources
4. Ensure that fish and other aquatic organism consumption **advisories are used by regulators as a short-term, temporary strategy** for informing those who consume and use fish, aquatic organisms and plants, and wildlife of risks while water quality standards are being attained and while prioritizing and pursuing the cleanup of contamination by appropriate parties; agencies must evaluate and address such risks, and require risk-producers to prevent, reduce, and clean up contamination of waters and aquatic ecosystems
5. Because many American Indian and Alaska Native (AI/AN) communities are particularly prone to environmental harm due to their dependence on subsistence fishing, hunting, and gathering, **conduct environmental research, fish consumption surveys, and**

monitoring, in consultation with federally recognized tribes and with the involvement of concerned tribal organizations, to determine the effects on, and ways to mitigate adverse effects on the health of AI/AN communities resulting from contaminated water sources and/or the food chain

6. Consistent with the 1988 EPA Indian Policy for the Administration of Environmental Programs on Indian Reservations, the federal trust responsibility to federally recognized tribes, and federal policies recognizing tribal sovereignty and promoting self-determination and self-sufficiency, **provide equitable funding and technical support** for tribal programs to protect AI/AN communities and tribal resources from harm caused by contaminated water and aquatic resources and, until tribes are able to assume responsibility for such programs, implement and require compliance with the federal environmental laws within Indian country; provided that, in consultation with tribes, EPA should promptly develop effective and appropriate regulatory strategies for setting, implementing, and attaining water quality standards within Indian country; and provided further that, EPA should work with Alaska Native villages to address the special circumstances that exist in Alaska and to protect the health of Alaska Natives from environmental threats associated with their extensive subsistence lifeways.

3 SURVEY APPROACHES AND SAMPLE SELECTION

3.1 Introduction

This chapter presents key issues to consider when selecting a survey approach and sample design for a fish consumption survey – Decisions about a survey approach and sample design. This is the third of five basic steps in the design and development of a survey for estimating fish consumption rates (as listed in Chapter 1). There is a need for budget-friendly, yet scientifically sound, approaches to estimate consumption rates and calculate risk associated with the consumption of chemically contaminated fish tissue (U.S. EPA, 1989). As states have increased their focus on this type of risk assessment, the need for site-specific fish consumption surveys has become more evident. In this chapter, we describe modes of data collection, sampling methods, sampling frames, and other considerations when selecting an approach to allow calculation of consumption rates for the population of interest. The discussions in this chapter assume the reader is familiar with elementary concepts in survey sampling; for an overview of survey sampling, see Lohr (2010).

It should be noted that the decision about a survey approach must be based on an overall assessment of a number of factors including, but not limited to, survey objectives, characteristics of the study population, the desired analytic approach, and budgetary factors. Sections 3.2 and 3.3 describe various types of surveys and sampling strategies that may be considered. At the end of this chapter, a series of tables organized by budget/funding levels (Tables 3-5 through 3-7) provide guidance as to when each approach may be the appropriate choice based on multiple independent factors such as the need for a representative sample, required response rates, characteristics of the population, prevalence of the behavior of interest, and others. These tables allow readers to take various study objectives and population characteristics into account within a budgetary framework.

If tribal populations are involved, consultation with the appropriate tribal governments is recommended in advance of the survey (see Section 3.3.2).

3.2 Types of Surveys

Approaches to conducting fish consumption surveys include single mode (e.g., telephone, web, mail, in-person interview, etc.), as well as mixed-mode (e.g., mail to phone, mail to web) designs. A discussion of each of these approaches is presented below. Most of these survey approaches can be used with any type of sampling frame, although some particular combinations may not be feasible due to budgetary constraints or population characteristics (see Tables 3-5 through 3-7).

3.2.1 In-Person Interview

In-person interviews can be conducted at known fishing locations, at the respondent's home, or at some other location (see, for example, CRITFC, 1994 and Michigan Department of Community Health, 2007). An in-person survey typically results in relatively high response rates because of the personal contacts made by interviewers, which may include some initial efforts to build the trust of

the respondent prior to the actual interview. For example, advance letters can be mailed to prospective survey participants to introduce the study and help establish credibility. In-person interviews allow for the collection of more complicated information because the interviewer can work with the respondent to obtain relevant information through a series of probes. Sometimes visual aids are used (e.g., maps, show cards with response categories, or pictures of fish species, portion sizes, etc.). Answers to survey questions can be recorded on hard-copy questionnaires or entered directly into a computer database on a laptop or tablet computer used by the interviewer. Interviewers are typically recruited for their ability to engage study participants, gain cooperation, and accurately record responses. They must be trained in the standardized data collection procedures developed for the survey, and informed about the overall goals of the research. Because of the opportunity for the interviewer to directly interact with the respondent, administer appropriate probes, establish time and location suitable for the respondent, and ensure that all questions are answered, in-person surveys typically result in the highest response rates and obtain high-quality data. They can, however, be quite expensive.

3.2.2 Creel Survey

A creel survey is a specialized form of personal interview to obtain information about fish that have been caught (see Zale et al., 2013). Creel surveys can be roving (meeting the fishers where they are fishing), or they can be stationary (meeting fishers where they return to the dock). In addition to asking a specific set of questions about fishing activity and fish consumption behavior, the researcher can try to identify and/or measure fish in the fisher's possession (the "creel"). The creel survey can be conducted at access points (e.g., boat ramps, docks), along the shoreline, or on the water from a boat. Fish consumption information obtained from the fishers is hypothetical since actual consumption has not yet occurred; however additional questions can be incorporated into a creel survey such as a 24-hour fish consumption recall.

A creel survey can also be used to develop a contact list of fishers who can later be sent a mail survey (or called for a telephone survey). This method provides confidence that actual fishers are being reached, rather than using random digit dial screening of households to locate fishers.

3.2.3 Telephone Survey

Telephone surveys contact selected respondents by telephone, using either their landline phone or their cell phone (for example, see Imm et al., 2005). Depending on the survey design, a brief telephone interview may be conducted to screen participants before inviting them to complete a more extensive telephone interview. Respondent answers may be entered either onto a hard-copy questionnaire or directly into a computer database by interviewers trained on the study content and procedures. To successfully reach a representative sample, a telephone survey should include cell telephone numbers as well as landline numbers.

3.2.4 Mail Survey

Mail surveys consist of a self-administered questionnaire that is mailed to respondents (see Minnesota Department of Health, 2012). Mail surveys are an inexpensive means for covering a larger or widely dispersed population of interest. The development of address-based samples has also allowed mail surveys to serve as effective screening tools for identifying subpopulations (Brick et al., 2012). Address-based samples can also be used to survey very small, targeted geographic areas.

Because they are self-administered, mail surveys lack the advantages offered by computerization of the instrument (available with in-person, telephone, or web/mobile modes), which can accurately navigate the respondent or the interviewer through the survey instrument and conduct real-time editing of entered data.

3.2.5 Web or Mobile Survey

Web surveys, like mail surveys, are self-administered survey tools that can collect recent and past fishing or consumption activities. However, web surveys offer the advantage of computerization, which allows for a more complex survey instrument to be administered. Web surveys may also take advantage of graphics such as pictures of different species of fish. Selected individuals are generally mailed information for accessing the web-based questionnaire (including a personal identification number, or PIN), or sent an email with a URL linking to the survey instrument. A limitation is that not all members of the target population may have web access.

Mobile surveys are similar to web surveys in that they are usually administered over the Internet, but can also utilize custom applications (or apps) that can be downloaded to a mobile device. Web surveys are often optimized for use with mobile devices, since for some populations a mobile device may be their primary access to the Internet (Pew Internet and Life Survey, 2013). Mobile surveys can be designed to collect recent and past consumption or fishing activities, but can also be used by in-person interviewers to capture other information—for example, photographs of fish, or to map locations where fish were caught.

3.2.6 Mixed Mode Survey

Mixed mode surveys take advantage of multiple survey modes in order to maximize response to the survey and, in many cases, minimize costs. There are a number of approaches to combining different data collection modes. One example that minimizes costs and maximizes response is to conduct a mail survey to collect fish consumption data of one (or more) household members. After the mail survey protocol has been completed (i.e., all nonresponse mailings have been completed), nonrespondents to the mail survey are contacted by interviewers to collect consumption information through a telephone interview. Another example that can minimize costs is to begin with a telephone mode. After the telephone contact protocol has been fulfilled (i.e., a specified number of contact attempts across times and days), nonrespondents are visited by an interviewer to collect fish consumption information for a selected person or household. A common mixed mode survey for dietary data collection with the goal of collecting multiple 24-hour recalls is to do the first 24-hour

recall in person and the follow-up 24-hour recalls by telephone. For more on mixed-mode survey designs, the reader is directed to de Leeuw (2005).

When utilizing mixed mode protocols, it is common practice to begin with the most inexpensive modes, proceeding to the more expensive modes in a sequential order. The researcher should be cautioned against offering respondents a choice of survey modes simultaneously (concurrent mixed mode protocols). Research has shown that this can lead to reductions in overall response compared to single mode designs (see Millar & Dillman, 2011, for an example with web and mail modes).

3.3 Types of Samples

A sample is typically defined as a subset of a larger population or group and through careful design, selection, and analysis, survey results from the sample can be representative of the population of interest. General population surveys, or probability surveys, are distinguished from other types of surveys by the fact that each unit (e.g., individual, household, etc.) in the population is given a known probability of selection that allows statistical inference to be used to generalize the survey results to the larger population – something that cannot be done with nonprobability samples. The first step in defining a sample is to decide whether the sampling unit will be households, individuals, events, or something else. For fish consumption surveys, the sampling unit is typically the individual consumer. When sampling (rather than taking a census of) the population, it would be inappropriate to consider all members of a household in a particular subgroup (e.g., children) as independent observations of the population because of obvious "household effects." If each individual in every household were considered an independent case, the consumption estimates for the population would be skewed toward those of larger families. If the individual is the sampling unit, an appropriate design might be to randomly select a household and then randomly select a household member within the target population.

With some sample designs (e.g., address-based sampling, telephone sampling, and area probability sampling), the ultimate unit within the sampling frame is the household rather than the individual. If the unit of analysis for a particular fish consumption survey is the individual, and one of these sampling frames is to be used, an additional stage of sampling is required. That is, after the sampled household has been identified, individuals must be sampled from among the eligible persons in the household. There are a number of options for choosing a method for within-household selection. Further details may be found in Gaziano (2005), which provides comparisons of various approaches for selecting survey respondents from among all household members.

As mentioned previously, the type of sample must be determined early in the design process, as many aspects of the sample design depend on the specification of the target population. For general population surveys (i.e., those in which the sample population is to be representative of—or generalizable to—for example, all adults, or all persons age 5 and older), the considerations and design choices are different from surveys that focus only on specific population subgroups (e.g., high consumers, women of child-bearing age, recreational fishers, etc.). General population surveys, or probability surveys, are distinguished from other types of surveys by the fact that each unit (e.g., individual, household, etc.) in the population is given a known probability of selection that allows statistical inference to be used to generalize the survey results to the larger population—something that cannot be done with nonprobability samples.

In general population surveys, nearly every household is expected to contain at least one member of the group of interest. Although there may also be the need to develop subgroup estimates (e.g., estimates by gender or age group), these goals should be balanced against the goal of obtaining estimates for the population as a whole. In contrast, when the survey aims to study only targeted populations (e.g., women of childbearing age), some proportion of households will not contain a member of this target population. These surveys typically involve a significant screening effort to determine whether households contain members of the target population. Screening for subgroups increases cost and requires a longer period for data collection.

In addition to defining the scope of the target population, another important aspect that characterizes the target population is the geographic area of interest (e.g., regional vs. a particular state vs. coastal counties in a particular state). The geographic area of interest may be defined, for example, by a need to assess consumption of fish caught in certain bodies of water.

3.3.1 Choosing a Sampling Frame

Depending on the target population, as well as overall survey objectives, there are a variety of possible sampling frames to choose from, each with their own strengths and weaknesses. Below, each of these is discussed in turn. In many cases, the choice of data collection mode will influence the type of sampling frame that is used. The quality of the sampling frame will depend upon the statistical methods used to select the sample, and the coverage, i.e., the degree to which the target population is covered by, or included in, the sampling frame. Researchers with limited resources should ensure that they choose the highest quality sample to meet their goals. It is important to choose a sampling frame that allows for a sample that is representative, while also employing a survey methodology that minimizes potential biases.

Procedures for obtaining and evaluating sample frames should be developed as a part of the study design. For telephone samples, address-based samples, or area probability samples, this will generally be an empirical question. One important aspect of the quality of these frames is coverage; i.e., the degree to which the target population is covered by, or included in, the sampling frame. All of the sampling frames described below, if used properly, can be used to generate probability samples. With appropriate use of sampling weights during the analysis, they can be representative of a larger population, thus providing results that are generalizable to that population (see discussion of weights in Chapter 7).

"Dual-frame" samples are developed by combining two of the sample types described in this section. A dual-frame sample is used when one type of sampling frame offers superior coverage of the target population but for which sampling and/or data collection are expensive, while another type of frame offers inferior coverage of the target population but at a lesser cost. Dual-frame sample designs, and the computation of estimates from these, can be complex, and requires careful consideration of the following:

- Allocation of the sample between the two frames

- De-duplication between the two frames (i.e., persons included on both frames), either prior to sampling, through identification of persons in the "overlap," or by screening out persons in the overlap sampled through one of the frames
- Any differences in data collection methods/procedures for the two frames
- Combining estimates from the two frames (which involves consideration of differences in response rates and coverage between the two frames)

Despite these complexities, dual-frame samples may serve particular populations well. For general population fish consumption surveys, for example, considerable efficiencies may be gained by using one sampling frame for a specific subpopulation (e.g., lists of licensed fishers if lists are available that truly represent the subpopulation of interest), combined with a broader sampling frame (e.g., address-based sampling) for the remainder of the population. This type of sample will likely require the expertise of a trained sampling statistician to successfully implement.

Random Digit Dialing (RDD) Sample and Dual-Frame (RDD with Cell Phone Sample)

A random digit dialing (RDD) sample is a sample of telephone numbers. Historically, RDD samples have been samples of listed and non-listed landline telephone numbers. Every working number in the population has an equal probability of selection. However, as the proportion of households having only cell phones has increased (estimated to be 43 percent during the first half of 2014; see Blumberg and Luke, 2014), survey researchers have increasingly turned to methods that include cell phone numbers as well as landlines in the sample.

Using both landline and cell phone numbers is a common example of a "dual-frame" sample (as generally described above). While the dual-frame telephone approach addresses the undercoverage associated with landline RDD sampling frames, it may exacerbate the problem of declining response rates to telephone surveys, as response rates for cell-phone components are generally lower than their landline counterparts (Brick et al., 2007; Link et al., 2007).

If a dual-frame RDD sample is desired, households with both landline and cell service may be sampled from either frame. There are two general approaches for accounting for a household's frame membership in dual-frame telephone surveys: (1) a screening approach in which landline households are screened out of the cell sample so that each household can, effectively, be included in the sample through only one frame; and (2) an overlap approach in which households are retained in the sample regardless of frame membership and dual-frame estimation methods are used. Unless a general population survey is required, the extensive screening required to identify target populations (e.g., high consumers of fish or pregnant or lactating women) is not feasible due to the high costs. Instead, the overlap approach can be used.

A key consideration is how to allocate the sample between the two frames. State- and county-level telephone service estimates are now available through at least one vendor, and these may be used to determine an efficient allocation among landline and cell numbers. State- and county-level telephone service estimates may be available from other vendors that provide telephone samples to survey organizations, and these may also be used.

Another consideration with dual-frame sampling is that while vendors maintain information that allows telephone exchanges (i.e., the first six digits of a telephone number) to be linked to

geographic areas, these linkages are known to involve error, so for surveys with geographically restricted target populations, it is necessary to confirm place of residence prior to completing the interview. To minimize bias from excluding people who have moved into an area while retaining their previous phone number, it is important to work with vendors that can identify some of these "foreign" numbers for inclusion in the sampling frame.

Pros/Cons: An advantage of a telephone sample is that it can cover a large geographic area and can be representative of the population (with telephones). However, it may be biased if a significant percentage of the population does not have a telephone. Due to the limitations of high cost (to identify sufficient numbers of subgroup members) and expected low response rates (due to telephone contacts), in general for fish consumption surveys, the use of RDD samples may be appropriate only when leveraging existing survey efforts such as the Behavioral Risk Factor Surveillance System, (BRFSS) (see discussion of leveraging in Section 3.12).

Intercept Sample

If properly implemented, an intercept sample (such as an in-person creel survey or an on-the-bank survey) can produce a probability sample when the survey population is restricted to people who must visit one of a limited set of locations. Such a design is venue- and time-based, and could be used to sample persons known to frequent certain types of locations; for example, to sample fishers who fish in a specific river at certain points in time. To ensure good coverage of the target population, intercept surveys should be conducted on different dates, at different times of day, and at different locations. Locations (often, open-access points) are identified, and for each location to be visited, survey days/times are randomly assigned and data collection (in the form of interviews with fishers) occurs during those days/times. At a given location during the sampled data collection period, there may or may not be sampling of persons. Often, a "take all" approach is used; in other words, every fisher accessing the body of water at the given location is sampled. Intercept surveys may also contain a separate data collection effort that counts fishers using the body of water. These counts are used in estimation to attribute proper coverage to the locations and the survey days/times. If subjects are to be interviewed more than once, contact information can be collected to allow for subsequent interviews to be conducted by mail or telephone.

An intercept survey alone would not be a desirable approach for a general population survey (e.g., a population that includes fishers and non-fishers) because of its very limited coverage. However, it may be a useful approach for sampling licensed and unlicensed fishers who fish in a specific body of water. Additional information to be obtained would include whether or not the fishers and/or anybody else, such as their family, actually consume the fish caught, or if they practice "catch and release" only. For further discussion of methods and potential challenges in conducting creel surveys, see Kinnell et al., 2007; Ray, Craven, Kinnell, Bingham, Freeman, & Finley, 2007; and Ray, Craven, Bingham, Kinnell, Hastings, & Finley, 2007. Also see Price, Su, & Gray, 1994.

Pros/Cons: The advantages of this design are that response rates are expected to be higher than telephone or self-administered modes because of the in-person contact. However, if interviewers are not fluent in languages/customs of ethnic groups likely to be encountered, or if interviewers are perceived to be associated with fish and wildlife or law enforcement groups, there may be an impact on response rates. Individuals interested in fishing may not be interested in providing information. Creel surveys allow for a high degree of flexibility in the use of visual aids, and screening costs are

low. Other disadvantages include difficulties in making the results representative of a larger population, and the high potential cost of covering large geographic areas. Additionally, consumption rate information gathered from a traditional creel survey may be hypothetical because consumption of the current catch has not yet occurred and the fish may be consumed by persons other than the fisher. Conducting a 24-hour fish consumption recall during the creel survey could eliminate this issue.

List Sample

In some cases, it may be possible to identify a list that contains a significant portion of the target population. For example, if recreational fishers are the target population of interest, lists of persons issued fishing licenses might be expected to cover a significant portion of the population. Similarly, a list of tribal members, or health registry lists (e.g., birth records), may serve as fairly complete listings for those subpopulations (Connelly et al., 2013). In these cases, such lists could serve as the sampling frame alone (with recognition of the implications of undercoverage) or could be used in combination with another approach that yields greater coverage (e.g., area probability sampling). However, access to this information may be limited to tribal members. Tribes may not participate in a survey conducted by an outside entity unless information is protected and the tribe has assurances that the objectives of the survey are in their best interest. Tribes may also wish to have control and input on the custody of data and analysis and interpretation of results. These issues will likely require tribal government approval of the use of tribal enrollment lists.

The information available on each list may influence or even dictate the data collection approach. It will be important to review what information is available for the list frame of interest before deciding on a data collection approach. Pilot testing (e.g., checking the completeness and/or validity of the list) may be helpful prior to using the list for sampling purposes. For example, if a particular list frame includes addresses for the entire frame, but telephone numbers for only a portion of the frame, then data collection methods utilizing telephone contact may not be easy to implement. Examples of useful information are provided below:

- Name
- Address
- Telephone number
- License type (fish/water type; lifetime/temporary/seasonal/yearly; other)
- Age
- Date license was obtained (can help determine whether information is current)
- For tribal enrollment lists, date each person was added to the list or when the list was last updated, as this is a possible indicator of whether information is current

With list sampling frames, there are list maintenance issues to consider. A common issue involves duplicates. For example, if a list of persons issued fishing licenses is used but a survey will be done at the household level, it may be necessary to remove duplicates. Another example of a situation in which de-duplication is necessary is one in which a license frame includes duplicate listings for persons issued more than one license (e.g., if licenses are issued for different bodies of water, and a single fisher holds multiple licenses). Also, if the list is actually a compilation of lists from different sources, it may be necessary to both remove duplicates and reformat the lists to achieve a consistent

format. In addition, steps should be taken to ensure that the most recent and up-to-date lists are used for the sampling frame.

Pros/Cons: The main advantage of a list sample is the potential for easy targeting of the population of interest. Disadvantages include the possibility of undercoverage because some people may fish without a license and the difficulty of developing high quality, unduplicated, and up-to-date lists.

Area Probability Sample

An area probability sample uses multi-stage sampling, with geographically based frames at each stage. The first stage involves the creation of primary sampling units (PSUs); for national surveys, PSUs are often counties or groups of geographically contiguous counties. The second stage involves the creation of secondary sampling units (SSUs) within sampled PSUs; for national surveys, SSUs are often census tracts or groups of census blocks. Within sampled SSUs, lists of households are compiled, and households are sampled in the third stage. Typically, the use of an area probability sample requires "boots on the ground" so that data collectors (i.e., listers) can compile lists of dwelling units (households) within a selected area. Quality concerns arise when listers miss dwelling units. This can be due to several factors, such as new areas of construction, hidden dwelling units (e.g., behind another dwelling unit, or above a business), or unidentified multi-unit structures (e.g., a house where basement residents are a separate household unit). This may be followed by sampling persons within households, which usually requires an in-person visit to the household. Although the stages discussed above describe a "typical" area probability sample, more or fewer stages of selection may be warranted, depending on the target population.

Since it results in a geographically clustered sample, area probability sampling is generally used when data collection is planned to be in-person. For example, if the plan is to administer the interview in-person or to use in-person interviewing to follow up nonrespondents, having a geographically clustered sample may be necessary from a cost and field management perspective.

Pros/Cons: The advantages of using an area probability sample are that it has the highest coverage of the population; thus the survey could be designed to be representative of the area sampled and it has the highest expected response rate. Disadvantages include the high cost and the probable complexity of sample clustering. Area probability sampling is relatively expensive, due to the need for in-person data collection, including household screening for the target population.

Address-Based Sample (ABS)

An address-based sample (ABS) is selected from sampling frames maintained by vendors that originate from the U.S. Postal Service's Delivery Sequence File (DSF) or Computerized Delivery Sequence (CDS) file. ABS frames may be constructed within areas as the last (or penultimate) stage of selection in an area probability sample (discussed above), or they may stand alone as a frame for a single-stage sample. The former approach may be used if it is necessary to cluster the sample within relatively small geographic areas; the latter approach may be sufficient if such clustering is not necessary.

Address-based samples rely on lists of addresses maintained by vendors. Providers of these lists generally strive for high-quality lists, but coverage issues can arise in certain areas. In the case of mail survey designs, households that rely on P.O. boxes for mail will be missed. If address-based samples are used for telephone approaches that match a telephone number to the sampling frame, the number matched may not always be to the address sampled. It will be necessary to verify the respondent's address to ensure the sampled address was contacted.

Because ABS sampling frames consist of mailing addresses, this type of frame lends itself well to a data collection approach that uses mail for the initial (or sole) contact with the household. However, if there are concerns about literacy among the target population, they can be addressed in the ABS design by using a multi-mode data collection approach or hybrid approaches (e.g., a design that involves a very simple mail screener followed by an in-person interview). With these variations, mail can be used to screen and identify the target population, or the screening could be done by telephone (using telephone numbers matched to addresses, which can generally be obtained for about 40 to 60 percent of addresses) or in-person. The screening survey may be completed by any adult household member, reducing the impact of low literacy levels. The substantive survey can then be administered either by telephone or in-person. A mail screener administered to an ABS sample will yield greater coverage than a telephone screener administered to an ABS sample (due to the inability to obtain accurate telephone numbers for some proportion of sampled addresses), at much lower cost than in-person screening. However, it may still be necessary to include telephone and/or in-person screening for nonresponse followup to the mail screener.

Pros/Cons: Advantages of an ABS are high coverage, moderate to high response rates (depending on mode of collection), saliency of the topic for those surveyed, and the ability to cover a large geographic area and be representative. Disadvantages may include high screening costs to reach a potentially hard-to-find target population (e.g., high fish consumers) and the potential for undercoverage in some rural areas. A survey using an ABS frame could be conducted through mail, a mail-to-web hybrid, or web only (with a mail invitation). An in-person or mail-to-phone hybrid ABS survey would be more expensive.

3.3.2 Challenges Associated with Some Sampling Frames

High Fish Consumers

If the goal is to estimate fish consumption rates for high fish consumers, it is necessary to take advantage of sampling frames that are likely to include high concentrations of this population. State lists of fishing license holders are common frames for consumption surveys since these are likely to target high fish consumers (Ashford et al., 2009). For subpopulations or targeted groups, such as tribes, there may be tribal lists or tribal enrollment records that could be used for sampling (with the proper approvals). The first step is to determine what lists are available and how access can be obtained. For some states, lists of license holders are publicly available, while for others, access may only be allowable through a state agency (e.g., fishery department, or fish and game agency). In some cases, state legislatures may restrict the release of personal information from license registries.

Many states have multiple licenses and requirements for who needs a license. Also, some states have "sportsmen" licenses that cover both hunting and fishing—further complicating the identification of fishers because avid hunters may only be occasional (or non-) fishers. There may be separate licenses for fresh water fishing, saltwater fishing, and/or shellfish. Further, there may be a primary and secondary license (e.g., conservation and fishing license; or fishing license and gear permit), and it may not be necessary to be a holder of both. For fish consumption surveys, it may be important to sample from multiple license lists, as consumption may vary by license type (Katner et al., 2011). In such instances, procedures to account for duplication across lists should be developed. Finally, license requirements may vary by residency or age. For example, senior citizens or tribal members may be exempt from license requirements or may not be required to purchase all types of licenses. Knowing this will help inform how well the license frame available will represent the population of interest.

Small, Isolated Populations and/or Targeted Geographic Areas

As discussed previously, surveys of smaller and/or more isolated populations present special challenges, including those discussed in Section 2.3.2.

If a specific area or body of water is a targeted area of interest, license requirements for that area or body of water should be reviewed. For some areas or bodies of water, there may not be a requirement to be a license holder and other frames or approaches may be necessary (e.g., see Kinnell et al., 2007).

Access to tribal lists or enrollment records may require cooperation and approval from the Tribal Council for the tribe (or tribes) of interest. If there are multiple tribes, cooperation and approval will be needed for each participating tribe. Tribes may have study review panels whose approval is needed before a Council will consider a study. It is important to understand the tribal approval process for each tribe whose participation is desired and the likely timelines for approval early in the planning stages of a study.

For larger geographic areas, approaches not requiring face-to-face contact (e.g., phone and mail surveys, diaries) could be more appropriate. Available staffing and time resources are also important considerations in selecting the survey approach since, for example, multiple interviewers can cover larger geographic areas simultaneously.

Characteristics of the Source of the Fish

The decision about which survey approach to use can depend on the source of the fish, certain characteristics of the fish, and how they are harvested. If the survey is mainly interested in self-caught fish, three important characteristics of fish sources are: (1) the number of access points; (2) the fishing pressure (i.e., the amount of fishing activity); and (3) the size of the geographic area. Access points refer to fishing locations for shore fishers (e.g., beach, riverbank, boat dock, fishing pier, etc.) and boat ramps for offshore fishers, as well as parking lots or preserve entrances where fishers might begin their activities. In locations with multiple access points, off-site approaches such as telephone surveys, mail surveys, and/or diaries may be preferred. An exception to this is the

roving creel survey, an on-site approach that can also yield good results in fishing areas with many access points, although a creel survey may not be representative.

In fishing areas with high fishing pressure, mail surveys, personal interviews, and access-point creel surveys may be effective because fishers are concentrated in relatively small areas. Roving creel surveys, in which the interviewer moves from fisher to fisher and sometimes from site to site, are more applicable to areas with low fishing pressure, where ample time is available for instantaneous counts and for interviewing all fishers.

If the survey objectives involve all sources of fish (including purchased fish and/or fish consumed in restaurants), the researcher must take this into consideration in the selection of the sample design and survey approach.

Seasonality

Ideally, a survey design would include observations throughout the year to capture information on practices across all seasons. There are several different ways to accomplish this. One way is to design a longitudinal survey to obtain consumption data at least four times in a year (i.e., once each season). This can be a difficult design to implement, as maintaining contact with the same respondents throughout the year could be costly, requiring rigorous case management and tracing and, ultimately, may result in too much sample loss to meet survey objectives. Another way to get data throughout the year is to divide the sample into smaller groups and begin the survey at different points throughout the year, as discussed in Section 2.3.3. This would provide snapshots of consumption patterns throughout the year. For the analysis, seasonality can be accounted for analytically by fitting terms for seasonal trends, although care must be taken to ensure sufficient sample sizes to cover each season.

3.4 Benchmarking

For general population surveys (and surveys of some specific subpopulations), it may be useful to benchmark the sampled population to external totals (e.g., population totals from the decennial U.S. Census or the American Community Survey). To the extent that such data are available from other studies for the subgroups of interest, such benchmarking can shed light on potential biases (due, for example, to undercoverage or differential nonresponse among subgroups), and may be used in adjusting for those biases. For example, a state survey might consider low-income women to be of particular interest. In this case, it would be useful to know how closely the percentage of low-income respondents matches the state-wide percentage of low-income women in order to be able to assess representativeness. To facilitate such benchmarking, it is important to plan for it in the survey design phase, to ensure that appropriate information is captured in the correct format (e.g., in a manner consistent with the way the information was captured in the external source). Benchmarking is also referred to as "post-stratification" or "raking." Additional information on benchmarking can be found in Holt and Smith, 1979, Oh and Scheuren, 1983, and Zhang, 2000.

3.5 Determining Required Sample Sizes

The sample size is the number of individuals selected from the sampling frame to be contacted for the survey. Decisions about sample size ultimately influence the analytical power of the survey and must take into account some assumptions about expected response rates. The term "respondents" refers to the people who complete the data collection and for whom data are available for analysis. The difference between the sample size and the number of respondents is determined by the response rate and rate of being in-scope (e.g., if the survey is of tribal members, what percent of people living in the area will be members of the tribe).

The target sample size depends on many factors including how individuals are selected, how each person's data are collected and analyzed, what is to be estimated from the data, and the desired precision of those estimates. Precision requirements are generally specified as the largest acceptable standard error or coefficient of variation (CV) of the estimate in terms of the ability to detect differences between subgroups or between estimates over time, or, simply, the ability to report a single estimated consumption rate with adequate precision. When specifying the precision requirements, it is important to consider which subgroups (if any) the sample should be designed to support. That is, do the specified requirements apply only to the sample as a whole, or do they also apply to particular subgroups? Separate precision requirements may be specified for different estimates, such as for mean fish consumption and for the 90th percentile of fish consumption, and for the whole sample or specific subgroups. Then, a different sample size would be calculated for each precision requirement. Often, desired precision requirements are re-examined (and adjusted) after preliminary sample sizes have been computed (based on preliminary precision requirements) and assessed in light of the budget for the survey.

The detailed process of determining how to select individuals for the survey is referred to as the sampling design. In any survey, nonresponse can be a major issue. Some individuals may not want to complete the survey or may not be located. When selecting people for the survey, an adjustment must be made to account for these expected nonresponders. For example, if the required number of completed interviews is 1,000, but only 50 percent of those sampled actually complete a survey, then 2,000 individuals must be sampled and contacted to achieve the target number of completed surveys. Similarly, if some proportion of the sampled group is expected to be ineligible for the survey, this must also be taken into account when determining the sample size. For example, if the survey is interested only in women of a certain age, and this characteristic cannot be determined ahead of time from the sampling frame, a much larger number of individuals should be sampled to ensure that a sufficient number of completed interviews for the target group will ultimately be available for analysis.

If a list of the target individuals is available, individuals from the list can be randomly selected, contacted, and interviewed. This process is called simple random sampling. However, a list is often not available. Alternatively, individuals can be selected by first randomly selecting communities or zip codes, then randomly selecting small areas within communities or zip codes and interviewing all individuals in the selected small areas. This process does not require a master list and, for in-person interviews, may be more cost effective. This is an example of cluster sampling (the communities and small areas represent clusters of people). The estimates from cluster sampling are generally not as precise as estimates from a simple random sample. The relative precision is represented by a "design effect." When calculating the target sample size, the design effect is a factor that accounts for how

the respondents are selected. A design effect of 2.0 means that, compared to using simple random sampling, twice as many individuals need to be interviewed to get estimates that have the precision that would be obtained if simple random sampling had been possible and was used. Simple random sampling has a design effect of 1.0. Since simple random sampling is rarely possible, the following discussion assumes a design effect of 2.0 as an approximation of the design effect that might be found in fish consumption surveys. Section 3.5.1 provides additional guidance for specifying the design effect. For more detailed information on sample designs and design effects, see Lohr (2010), or contact a statistician.

Another consideration for specifying the target sample size is how the data, in this case fish consumption data, are collected and analyzed. Two basic approaches for obtaining usual intake rates are:

1. Contacting each person two or more times and, on each contact, collecting information on recent fish consumption
2. Contacting each person once and collecting as much information as possible on long-term fish consumption

The first approach is preferable from an analysis perspective because it reduces recall bias (respondents can more accurately remember recent, as opposed to long-term, fish consumption) and because it allows for correcting the estimates for random differences between interviews. However, since this approach requires more complicated statistical modeling and more respondent contacts, the second approach may be more feasible if resources are limited. Different analyses are required for these two data collection approaches, and the equations used to estimate the target sample size depend on which approach is used.

Finally, the target sample size depends on what is to be estimated and the desired precision of those estimates. The survey data might be used to estimate the mean of usual fish consumption across the population or the 90th percentile of usual fish consumption. The 90th percentile is the level of fish consumption that is greater than 90 percent of the target population and less than 10 percent of the population. Some parameters can be estimated more precisely than others. In most cases, the mean can be estimated with better precision than an extreme percentile, such as the 90th percentile. A 95 percent confidence interval is an interval that almost certainly (i.e., with 95% probability) covers the true values being estimated. A confidence interval for mean fish consumption will usually be narrower than a confidence interval for the 90th percentile. Increasing the sample size increases the precision of the estimates, resulting in shorter confidence intervals. As a result, estimating the mean to a desired precision may require fewer respondents than estimating the 90th percentile with the same precision.

Precision can be expressed in different ways. A confidence interval or a standard error is a measure of how precise an estimate is. The coefficient of variation is a measure of relative precision. For example, mean fish consumption might be 12 grams plus or minus 1.2 grams (1.2 is the standard error). Thus, the estimate of 12 is most likely (about two-thirds of the time) to be within 1.2 grams of the true value. The coefficient of variation is the ratio of the standard error to the mean. For this example the coefficient of variation is 1.2 divided by 12 or 0.10. This can also be expressed as 10 percent. Thus, the estimate of 12 is most likely within 10 percent of the true value. A 95 percent confidence interval might also be used. The half width of a 95 percent confidence interval is roughly twice the standard error. For example, the true fish consumption rate is also certainly (roughly

95 percent of the time) within the mean plus or minus two standard errors; e.g.; 12 +/- 2.4 or between 9.6 and 14.4. Similarly, the true value is almost certainly within 20 percent of the estimate of 12. In general, increasing the sample size by a factor of four reduces the standard error, the coefficient of variation, and the width of the confidence interval by a factor of two.

3.5.1 Design Effect

The design effect is an adjustment factor in the calculation of the sample size that is primarily determined by how the respondents are sampled. The design effect for a simple random sample from a list is 1.0. For other sample designs, the design effect will typically be greater than 1.0 (although stratified samples may yield design effects less than 1.0). Although design effects can be very large, for most practical applications involving well-designed samples, design effects will generally be less than 4.0. Detailed information for specifying the design effect is often not available. As a result, it may be necessary to consult a survey statistician for guidance and recognize that any estimate of the design effect is imprecise.

The best estimates of the design effect for the planned survey are obtained from historical surveys of the same (or a similar) population using the same (or very similar) sample design and collecting similar information. Thus, if a survey is conducted every three years using the same sample design, the sample size calculations for the second and later surveys can benefit from the estimated design effect from the earlier surveys. Software specifically designed for the analysis of survey data will generally provide estimates of the design effects. The design effect can be somewhat different for each parameter being estimated. Thus, if a previous survey does not have estimates of the parameter of interest, a fallback would be to use an average design effect across several related parameters.

Suppose, based on the design effects from an historical survey, it is decided to double the sample size while keeping the same sample design. Doubling the sample size while keeping the same design is equivalent to repeating the survey twice. For random samples from a list, this corresponds to randomly selecting twice as many list items (perhaps names or addresses) from the list. For cluster or area samples with PSUs (clusters) and SSUs (within cluster items), this corresponds to selecting twice as many clusters (PSUs) while retaining the historical within cluster sample selection procedures.

In general, we recommend contacting a survey statistician to evaluate the design information available from historical surveys and estimate an appropriate design effect for the planned survey. The statistician can use information from historical surveys and make adjustments for differences between these and the planned design to provide guidance on what design effect to use and other aspects of the survey design.

3.5.2 Sample Sizes for Multiple Contact Surveys

When the survey design requires contacting each person two or more times and, on each contact, collecting information on recent fish consumption, the sample size calculation requires an estimate of the expected probability of reported fish consumption on any one contact. Call this probability P.

Higher values of P result in smaller sample sizes. P depends on the sample design. For any particular type of fish consumption, higher values of P can be obtained by defining the surveyed population in a way that includes those with fish consumption and excludes, as much as possible, those without fish consumption. P using a 7-day recall will be higher than when using a 24-hour recall, though a 7-day recall results in data with more measurement error.

P can often be estimated from previous surveys of similar populations using similar data collection protocols. Since risks from fish contamination are particularly important for those with high fish consumption, using the proportion of the previous surveyed population that consumed fish may provide a conservative sample size estimate. Given different proportions of fish consumption in different demographic groups, the proportion for a demographic group with less frequent fish consumption might be used.

For example, using the 2007-2012 NHANES data, the proportion of adults reporting any fish consumption in a 24-hour recall is 0.18, or 18 percent, on average across the United States. This value varies among demographic groups and is smaller when looking at specific types of fish. For example, 13 percent of adults reported consuming freshwater or estuarine fish in the previous 24 hours. The following table shows the proportion of respondents in various demographic groups who reported fish consumption on a 24-hour recall based on the NHANES data. These values can provide a starting point for estimating P when other relevant data are not available. NHANES does not provide good estimates for finer categorization of race and ethnicity than is displayed below (though NHANES 2001-2012 does provide data for non-Hispanic Asians). If the target population is American Indian/Alaskan Native, for example, a more relevant source should be used.

Table 3-1. Estimated percentages of the U.S. population who consumed fish on a given day, NHANES 2007-2012

Demographic group	Type of fish consumed			
	All fish	Freshwater and estuarine	Freshwater	Marine
Children < 18 years	8.6%	5.1%	2.0%	7.9%
Adults (18 and over)	18.0%	13.0%	5.6%	16.4%
Mexican American Adults	15.8%	11.8%	5.4%	14.0%
Other Hispanic Adults	16.2%	11.2%	4.0%	15.5%
Non-Hispanic White Adults	16.9%	11.8%	4.9%	15.6%
Non-Hispanic Black Adults	19.0%	14.6%	6.8%	16.3%
Other Race Adults	33.5%	27.0%	12.6%	31.2%
Adults with household income between $20 and $45k	14.9%	10.0%	4.7%	13.5%
Adults with household income between $45 and $75k	17.4%	12.3%	5.2%	16.2%
Adults with household income over $75k	21.5%	16.5%	7.0%	19.7%

Analyzing the reported fish consumption data from multiple contacts (or recalls) requires estimating variance components: the variability in the amount of fish consumed by the same individual on different days and the additional variation between people (some people consume fish more often or usually consume more fish than others). Estimating the variance components requires that multiple respondents report some fish consumption in at least two recalls. The algorithm may not converge or the estimates of percentiles may be particularly imprecise if there are few respondents who reported fish consumption on at least two recalls. Let M be the number of respondents reporting fish consumption in at least two recalls. Based on simulated data, the probability that the NCI method fails to converge and provide an estimate of UFCR increases when M drops below 20. Considering this as a minimum criterion, using the expected probability of reported fish consumption in any contact (P), a minimum sample size (N_{Min}) can be calculated as follows:

$$N_{Min} > \frac{20}{\mathrm{P}^2} * Deff$$

where $Deff$ is the design effect.

Using this minimum sample size, the number of respondents with at least two reports of fish consumption will almost certainly be greater than 20. If the overall probability of consuming fish in any one contact is 10 percent, that is, on average fish is consumed once every 10 days using a 24-hour recall, and a design effect of 2.0 is assumed, then:

$$N_{Min} > \frac{20}{0.1^2} * 2.0 = 4{,}000.$$

The following table shows the minimum sample size versus the expected probability of consuming fish in any recall and the design effect. The minimum sample size increases significantly as the expected proportion of contacts with reported fish consumption decreases. If the minimum sample size is unacceptably large, consider increasing the length of the recall period or consider revisions to the sample design to increase the expected P.

Table 3-2. Minimum sample size versus the expected probability of consuming fish in any recall (P) and the design effect

Design Effect	P							
	0.02	0.05	0.10	0.15	0.20	0.30	0.40	0.50
1.0	50,000	8,000	2,000	889	500	223	125	80
1.5	75,000	12,000	3,000	1,334	750	334	188	120
2.0	100,000	16,000	4,000	1,778	1,000	445	250	160

The equations above specify a minimum sample size for reliably calculating the survey estimates of the usual fish consumption rate. That minimum sample size may not provide estimates with the desired precision. The following formula calculates the target sample size (N) required to obtain the desired precision of the UFCR estimate:

$$N_{CV} = \frac{A}{CV^2} * Deff$$

where $Deff$ is the design effect, CV is the desired coefficient of variation of the estimate, and A is a constant that depends on what parameter is to be estimated and on the expected probability of reported fish consumption.

The following table shows values for A when estimating the mean, the median (50th percentile), or the 90th percentile of usual fish consumption.

Table 3-3. Values for A when estimating the mean, the median (50th percentile), or the 90th percentile of usual fish consumption

Expected Probability of Fish Consumption (P)	Parameter being estimated		
	90th Percentile	Median	Mean
0.05	6.46	4.35	5.20
0.10	5.77	4.26	4.34
0.20	4.96	3.77	3.31
0.30	4.29	3.44	2.83
0.45	3.76	3.13	2.36
0.60	3.40	2.76	2.04
0.75	3.21	2.56	1.91
0.90	2.98	2.44	1.77
0.95	3.00	2.43	1.76

For estimating the 90th percentile of UFCR when the expected probability of fish consumption is 10 percent, the constant A is 5.77, from the second line in the table above. If the desired precision specifies a confidence interval of plus or minus 10 percent (corresponding to a CV of 5% or 0.05) and using a design effect of 2, the desired sample size would be:

$$N = \frac{5.77}{.05^2} * 2 = 4{,}616 \text{ respondents}$$

The final target analytical sample size (N) is the maximum of N_{CV} and N_{Min}.

$$N = \text{Max}(N_{Min}, N_{CV})$$

The constants in the table above were determined by simulating fish consumption data similar to the NHANES data with two 24-hour dietary recalls, analyzing the data using the NCI method (see Chapter 7) and calculating the precision of the estimates. The simulations used various assumptions representing a range of values that might be appropriate for 24-hour recall data. The values above are the 75th percentile estimates across the simulated conditions and thus may be conservative (i.e., high on average). However, the characteristics of other data collection approaches using other populations are unknown. Thus, these values are provided only as general guidance to sample size determination. Also, these target sample sizes must be increased to account for nonresponse (individuals who cannot be located, are not eligible for the survey, or who refuse to complete the survey).

If a 7-day fish recall is used for data collection, the probability of reporting fish consumption in any one contact (i.e., 7-day recall) will be higher than when using a 24-hour recall. The equations above can be used to approximate the target sample size if the probability of reporting fish consumption in any 7-day period can be estimated. The values of A from Table 3-3 are likely to provide conservative estimates of sample size when using a 7-day recall period.

3.5.3 Sample Sizes for Single Contact Surveys

If the second data collection approach is used (i.e., a fish-specific dietary assessment instrument, or fish diet screener, asking individuals to report long-term fish consumption amounts, or values from which long-term fish consumption can be calculated), another formula can be used to determine the optimal sample size.

If we have data on usual fish consumption from N individuals, we can calculate the proportion of the population with usual fish consumption above some specified value, such as 10 grams per day, and calculate a confidence interval for that proportion. If we specify the desired precision for the proportion estimate and the expected proportion (p), we can then calculate the sample size using the following formula:

$$N = \frac{3.84\, p(1-p)Deff}{H^2} < \frac{Deff}{H^2}$$

where H is the desired half width of the 95 percent confidence interval for the proportion. The ratio on the right above, provides an easily calculated, slightly conservative sample size estimate for when the expected proportion is unknown or close to 0.5. The table below shows the sample size as a function of the desired confidence interval width for proportions close to 50 percent and 90 percent. Notice that the confidence interval gets smaller (the precision increases) as N increases.

Table 3-4. Sample size as a function of the desired confidence interval width for proportions close to 50 percent and 90 percent

P	H	Confidence Interval	Design Effect		
			1	2	3
0.5	0.20	.30 to .70	24	48	72
0.5	0.10	.40 to .60	96	192	288
0.5	0.07	.43 to .57	196	392	588
0.5	0.05	.45 to .55	384	768	1152
0.5	0.04	.46 to .54	600	1200	1800
0.5	0.03	.47 to .53	1067	2134	3200
0.5	0.02	.48 to .52	2400	4800	7200
0.9	0.10	.80 to 1.00	35	70	104
0.9	0.07	.83 to .97	71	142	212
0.9	0.05	.85 to .95	139	277	415
0.9	0.04	.86 to .94	216	432	648
0.9	0.03	.87 to .93	384	768	1152
0.9	0.02	.88 to .92	864	1728	2592

Although the confidence interval above is not the same as a percentile, Table 3-4 can be used to calculate approximate sample sizes for percentiles. For example, when estimating the 90th percentile of reported long-term fish consumption, a sample size of 277 will provide an estimate of the 90th percentile that is almost certainly between the true 85th percentile and the true 95th percentile with a design effect of two.

3.6 Effect of Response Rates

Response rates affect the precision of the estimates derived from survey data. The potential for high response rates varies by mode of administration, with the most costly modes generally resulting in higher expected response. This expectation should be carefully considered when determining an appropriate sample size, particularly if the survey's objectives require the inclusion of a certain number of responses from subgroups (e.g., fish consumers).

If the response rate is expressed as the fraction of sampled respondents that provide complete data, the target sample size is:

$$Sample\ size = \frac{N}{Response\ rate}$$

All other things being equal, a survey with an expected response rate of 30 percent will require twice the potential sample size of a survey with an expected response rate of 60 percent. Lower response rates can result in more bias in the study findings as the likelihood of obtaining information on the target group of interest may be greatly diminished (e.g., high consumers). With very low response rates, additional biases may result from the few who are willing to participate being quite different from the population of interest in general. Thus, it is important to develop a realistic response rate

assumption. This may be based on other similar studies, although each survey has unique qualities that may affect response rates. These include data collection modes used, length of the questionnaire, type of respondent, type of questions asked, saliency of the subject matter to the respondents, experience of the interviewers, etc. (see Section 5.7.)

Within the current overall environment of declining response rates to surveys in general, the use of a well-designed monetary (or non-monetary) incentive program can be beneficial and help boost response rates if administered properly. The timing (e.g., incremental incentives or one-time), amount, and type of incentive should be considered with an understanding of the target population, the perceived burden of participation in the survey, and the study design. For further information on incentives and their effect on response rates, see Mercer et al., 2015. Additionally, when working with tribal communities, it would be very beneficial for those conducting the surveys to be familiar with the cultural protocols related to approaching an individual to share information. These protocols will vary by tribe, but if respected will result in improved response rates.

3.7 Accuracy

The required accuracy of the consumption rates is an important topic to be considered when establishing the survey objectives. All approaches outlined in this document can provide highly accurate estimates if resources are sufficient, valid survey designs including adequate numbers of respondents are used, and the design takes into account the characteristics of the subject matter and the target population.

There are several components to accuracy, including reliability (i.e., the variability of repeatability of the response); validity (the ability of the respondent to provide the correct answer); measurement errors (associated with the interviewer, the respondent, the questionnaire, and the data collection mode); bias (the overestimation or underestimation due to sampling or non-sampling errors); and random errors.

Other factors influencing the accuracy of the survey responses include whether the respondent views the survey as nonthreatening or sensitive; whether respondents remain anonymous; the length of the recall period (recall bias); whether some subgroups of the targeted population were excluded from being part of the survey (coverage bias); the tendency for respondents to provide responses that conform to ideal norms or what they believe the interviewer wants to hear, or that might enhance their self-image (prestige bias); the clarity of questions (question misinterpretation); the familiarity of the respondent with the subject matter; and the amount of specificity in the questions (e.g., exact numbers vs. approximations or ranges) (Wentland & Smith, 1993).

For fish consumption surveys focused on a limited subset of species and capture locations, the potential for misidentification of fish consumed is affected by recall bias, prestige bias, and the familiarity of the respondent with the subject matter. Some ethnic groups may have their own names for different species, may group some together under one common name, or there may be local nicknames for species or groups of species. This could be addressed by having the community be part of the survey design, and through pre-testing. The potential for misidentification is lowest for creel surveys and on-site in-person interviews (e.g., on-the-bank surveys) because the interviewer can both directly observe fish catch and/or allow respondents to visually select the species consumed

from displays of fish species. Survey participation affects the accuracy of consumption estimates by affecting the number and characteristics of respondents who are included in the survey. Also, with creel surveys, typically only the catch between the start of angling and the interviewer/respondent interaction is recorded, though as mentioned previously, a 24-hour fish consumption recall could be added to the traditional creel survey.

In general, surveys that allow all of the targeted population a chance to be included in the sample (e.g., do not exclude consumers living further away from a water body), include a larger number of respondents, and have a high response rate, provide a more accurate representative estimate of behaviors among the target population. An understanding of, and sensitivity to, the characteristics of the target population of concern can help to minimize nonresponse bias due to culture, religion, language, and attitudes toward government and authority. See Tarrant & Manfredo, 1993, and Vaske et al., 1996, for additional information.

3.8 Sample Stratification and Oversampling

Stratification of the sample is generally used for one (or more) of the following purposes: (1) to improve the precision of overall estimates; (2) to ensure adequate subgroup sample sizes; and (3) to improve the precision of subgroup estimates. Stratification involves partitioning the population into mutually exclusive, exhaustive groups (strata). Samples are generally selected independently within each stratum. See Lohr (2010) for further discussion of how strata should be formed and how they may be used to effectively achieve those objectives.

Stratification with proportional allocation (i.e., allocating the sample to the strata, proportional to the population in each stratum) may be used to achieve the first and second objectives. Achieving the third objective generally involves oversampling the rarer subgroups.

For fish consumption surveys, it may be desirable to stratify, for example, on the following:

1. Region/geographic area (e.g., proximity to a particular body of water), to support the production of estimates by area
2. Age and gender, to support consumption estimates for women of child-bearing age while limiting the sample sizes for other age/gender subgroups
3. Fishers vs. non-fishers, to ensure adequate sample sizes in each group for comparisons between these two groups
4. High consumers vs. others

3.9 Response Bias/Avidity

There are several types of response bias that are specific to fish consumption surveys. First, if fish consumption advisories have been issued, affected groups may be inclined to provide socially or culturally desirable responses to questions about whether or not they consumed fish (or certain types of fish). That is, respondents may be hesitant or may refuse to report that they consumed fish covered by an advisory. If there are restrictions on catching certain types of fish, respondents may

be reluctant to admit they did so, especially if the interviewer has a connection to some state or local authority. Survey designers must take care to introduce questions about fishing practices and/or consumption with an objective statement reassuring the respondent that there are no right or wrong answers to the questions. An assurance of confidentiality must also be provided.

Second, avidity bias is systematic error in estimates that results from avid fishers (who presumably are high consumers) having higher propensities to be selected and/or to respond to the survey (see Andrews et al., 2010, for discussion.) In order to estimate and correct for avidity bias, it is necessary to obtain angling frequency information as part of the data collection effort. If those who catch (and consume) more fish have different behaviors/characteristics than those who fish (and consume) less, it may be necessary to apply avidity weights during data analysis.

3.10 Choosing an Approach

Selecting a survey mode and sampling frame from among the options described above requires consideration of multiple factors, including the realities of possible budget/resource limitations. In this section, we present survey design options within the context of three different levels of funding/resources. These range from relatively low-cost approaches to more resource-intensive efforts. The more resource-intensive efforts are generally accepted as standard methodological approaches where the budget is not as tightly constrained. Each of the cost categories (presented in Tables 3-5, 3-6, and 3-7) offers methodologically sound choices for accomplishing a fish consumption survey. To help with decisions about investing adequate levels of resources needed to obtain high-quality, defensible data, survey developers may use decision-making tools such as Value of Information (Miinelli & Baio, 2015) or influence diagrams (Carriger & Barron, 2011).

Resource-sensitive options are presented using the following levels.

Level 1 (least expensive): Low- to moderate-cost partnerships that leverage existing surveys
Level 2: Low- to moderate-cost options for developing a new survey
Level 3 (most expensive): Moderate- to high-cost options for developing a new survey

Level 1: Low- to Moderate-Cost Partnerships that Leverage Existing Surveys

If resources are limited, a good approach would be for survey developers to first determine whether there may be a survey already in place within their geographic area of interest focusing on the population of interest. Any existing survey should be carefully evaluated to determine the feasibility, cost, and schedule for adding desired questions. This has the advantage of utilizing an existing survey methodology that has been established, without having to invest in developing an entire sample design and survey infrastructure. However, the frame and survey methodology of the existing survey should be carefully examined to determine whether or not it will result in a sample that is representative of the population of interest. Also this analysis should consider any potential biases or inaccuracies specific to the population of interest. For example, is seasonality addressed? Another consideration is whether or not the size of the existing effort is sufficient to support derivation of the desired statistics.

While there may be costs associated with adding questions to an existing survey, these costs should be less than what would be required to develop and implement a new survey. The process for adding questions onto existing surveys and cost for doing so will vary, depending on which survey is chosen as the best vehicle for obtaining the desired data. All of the level 1 survey options involve building on existing data collections.

There are several different types of survey designs that could be leveraged. One is a random digit dialing telephone survey (e.g., CDC's Behavioral Risk Factor Surveillance Survey, or BRFSS). BRFSS is conducted by the Centers for Disease Control and Prevention (www.cdc.gov/brfss) in all 50 states, the District of Columbia, and three U.S. territories. It is the Nation's leading telephone survey, collecting state data about U.S. residents regarding their health-related risk behaviors, chronic health conditions, and use of preventive services. BRFSS completes more than 400,000 adult interviews each year, making it the largest continuously conducted health survey system in the world. With technical and methodological assistance from CDC, state health departments use in-house interviewers or contract with telephone call centers or universities to administer the BRFSS surveys continuously throughout the year. States use a standardized core questionnaire, optional modules, and state-added questions. The survey uses random digit dialing (RDD) techniques for both landlines and cell phones.

While the content of the core BRFSS questionnaire is determined by state coordinators and the CDC, state coordinators may choose to add new questions based on submitted proposals (http://www.cdc.gov/brfss/about/brfss_faq.htm). There is a review process in place for BRFSS to decide whether additional questions can be accommodated and what fees will be charged. Before beginning this review process, the first step is for survey developers to determine whether a telephone survey is appropriate for their population of interest, and whether BRFSS would provide adequate coverage within the population(s) most of interest for the proposed surveys. There may also be limitations on the number of questions that can be added to the existing survey. Adding questions to BRFSS would yield fish consumption rates for the general population of the state. However, using BRFSS may not provide valid estimates for subpopulations of interest such as local fishers, pregnant women, or subsistence fishers.

Other types of surveys that could be leveraged include in-person intercept creel surveys, or web or mail surveys conducted within certain areas (i.e., surveys conducted by local agencies). For example, fisheries managers in a number of states have current or planned creel surveys that are done on a continuous basis, conducted by state, tribal, or local agency staff (examples include Florida FWC, 2014; Michigan DNR, 2014; Minnesota DNR, 2011; Oregon DFW, 2014; Texas PWD, 2014; Washington DFW, 2014). There are a variety of mail and online surveys that are also conducted by state and local authorities. Examples include the Tennessee Wildlife Resources Agency's (TWRA) BITE (Bass Information from Tournament Entries) program, which is a coordinated effort to obtain tournament data to the TWRA Fisheries Management Division via an online reporting form or a mail-in tournament report card. This information supplies information such as the number of participants, total catch, size, and weight structure of the tournament catch (Tennessee Wildlife Resources Agency, 2014). Also, the Florida Fish and Wildlife Conservation Commission created an online web survey to indicate where they harvest scallops, how many they collect, and how long it takes to harvest the shellfish (Florida Fish and Wildlife Conservation Commission, 2014). Also, New Jersey has undertaken the Recreational Saltwater Angler Survey to collect information on recreationally important marine finfish species (New Jersey Department of Environmental Protection, 2014). These are just a few examples of local surveys.

If it is decided that a survey's objectives can be fulfilled by leveraging an existing survey effort, the researchers should conduct a thorough assessment of the current possibilities that exist within the locales of interest. Then, negotiations would need to be held with the agency that conducts the survey to assess the cost and feasibility of adding the desired fish consumption questions to an existing survey.

Table 3-5 shows how each of these modes and sample designs could address survey objectives, and the pros/cons of each approach within the overall structure of a partnership with an existing survey.

Level 2: Low- to Moderate-Cost Options for Developing a New Survey

At this level, there are four modes of administration to choose from (mail only, mail-to-web hybrid, web only, or an in-person survey), and three different sample designs (address-based sampling, list sample, and an intercept design for the in-person mode). Under Level 2, these would be stand-alone new survey efforts, and all of the designs rely on relatively low-cost methodologies that can provide statistically sound results.

Table 3-6 shows how each of these modes and sample designs could address survey objectives, and the pros/cons of each approach. The mail-to-web option mode involves sending a mail invitation to potential study participants, inviting them to visit a website to complete the survey. If necessary, some screening can occur at the time of the mail contact to allow the survey to be targeted toward respondents with characteristics of interest (e.g., fishers).

Level 3: Moderate- to High-Cost Options for Developing a New Survey

This highest level presents some other generally accepted survey approaches that require more extensive resources to develop and implement. These are presented in this document to provide an understanding of what these approaches can offer compared to the low and moderate cost options, and when they may be most appropriate to consider. Three administration modes are within this level: telephone surveys, in-person surveys, and mail-to-telephone hybrid. Within each mode of administration, there are different sampling approaches that could be employed, depending on the characteristics of the population under study and other survey objectives. Table 3-7 shows how each of these modes and sample designs could address survey objectives, and the pros/cons of each approach.

In sum, Tables 3-5 through 3-7 describe major factors affecting the decision about an appropriate sample design and mode of data collection within these three budgetary frameworks. Each table lists independent survey objectives that the designer may wish to address, along with potential characteristics of the population to be surveyed. These tables are designed to aid the reader in choosing an appropriate sample design and data collection mode based on multiple independent factors, considering the known level of resources available for the planned survey. Listed on each table are the key factors to consider in choosing an approach, as follows.

Required response rates: Achieving the highest possible response rates is optimal. Higher response rates generally increase the precision of the estimate and maximize the number of observations available for analysis.

Need for use of visuals to aid respondent: Some modes of administration may not be suitable if visual aids are needed to obtain accurate information from the respondent. For example, for fish consumption surveys, visual aids showing portion sizes, species of fish, and/or maps of fishing locations may be helpful in improving the accuracy of responses.

Need for data over multiple seasons or points in time: Fish consumption rates can be heavily influenced by the time of year, and the design of consumption surveys should take into account when the survey will be administered. To estimate UFCR, interviews should take place at multiple points in time.

Size of geographic area to be covered: Some data collection modes and sampling frames may be better suited to smaller or larger geographic areas. For example, an in-person interview that needs to cover a large geographic area will be resource intensive due to the need for interviewers to move throughout the area.

Expected prevalence of the characteristic and/or behavior among the target population: What percentage of the population within the defined geographic area is likely to possess the demographic characteristic(s) of interest or behavior (e.g., pregnant women, children, elderly persons, high fish consumers, etc.)? If the characteristic or behavior is expected to be relatively rare among the population, then a larger sample size will be required so that completed interviews are obtained from a sufficient number of individuals with the characteristic and/or behavior of interest.

Literacy rates among population of interest: Certain modes of administration (e.g., any self-administered survey) may not be feasible for populations that are known to have low literacy rates.

Fixed addresses within population of interest: Certain modes of administration (e.g., a mail survey) are not feasible for highly mobile populations (e.g., fishers who may move with the seasons) because they will not receive the survey in the mail in a timely manner.

Web penetration rates within population of interest: While proven to be an efficient mode of data collection among certain groups, web surveys should not be planned for a population that is known to have limited Internet accessibility.

Telephone coverage rates for population of interest: Telephone surveys are not generally used for a population without a known telephone number (e.g., list survey) or availability of a telephone (e.g., RDD survey). Ideally, both landline and cell phone numbers will be available.

Length of the interview: Some modes of administration are not generally used for longer survey instruments. In general, the shorter the interview, the more options the researcher has for choosing a mode.

Table 3-5. Level 1 framework for choosing a sample design and data collection mode based on independent factors

YES = APPROPRIATE MAYBE = MAY BE APPROPRIATE NO = NOT APPROPRIATE	LEVEL 1: POTENTIAL LOW TO MODERATE COST PARTNERSHIPS TO LEVERAGE EXISTING SURVEYS		
	RDD Telephone Survey	In-Person Intercept Survey	Web or Mail Survey
SURVEY OBJECTIVES			
What response rates are needed from the survey to meet precision requirements?			
High (>80%)	NO	MAYBE	NO
Moderate (60-79%)	NO	MAYBE	MAYBE
Low (<60%)	YES	YES	YES
Are visuals needed for validity (e.g., river maps, pictures of fish species, portion sizes, etc.)?			
Yes	MAYBE	YES	YES
No	YES	YES	YES
Are data needed over multiple seasons, or are multiple observation points needed?			
Yes	YES	MAYBE	YES
No	YES	YES	YES
What is the size of the geographic area to be covered?			
Large	YES	NO	YES
Moderate	YES	MAYBE	YES
Small	YES	YES	YES
POPULATION CHARACTERISTICS			
Prevalence of characteristic or behavior of interest among target population (e.g., pregnant women and/or high fish consumers)?			
High percent of population	YES	YES	YES
Moderate percent of population	MAYBE	MAYBE	YES
Low percent of population	NO	MAYBE	YES
Unknown	MAYBE	MAYBE	YES
What are the literacy rates among the target population?			
High	YES	YES	YES
Moderate	YES	YES	MAYBE
Low	YES	YES	NO
Unknown	YES	YES	NO
How important is it to have high coverage of persons with no fixed address?			
Very important	NO	YES	NO
Moderately important	MAYBE	YES	MAYBE
Not important	YES	YES	YES

DRAFT

Table 3-5. Level 1 framework for choosing a sample design and data collection mode based on independent factors (continued)

	RDD Telephone Survey	In-Person Intercept Survey	Web or Mail Survey
What are the web penetration rates for the target population? *			
High	**YES**	**YES**	**YES**
Moderate	**YES**	**YES**	**YES**
Low	**YES**	**YES**	**WEB MAYBE** **MAIL YES**
Unknown	**YES**	**YES**	**WEB NO** **MAIL YES**
What are the telephone coverage rates for the target population? *			
High	**YES**	**YES**	**YES**
Moderate	**MAYBE**	**YES**	**YES**
Low	**NO**	**YES**	**YES**
Unknown	**NO**	**YES**	**YES**
PROS/CONS OF EACH APPROACH			
Coverage	Without dual frame, would exclude HHs w/o landline telephone numbers. In mid-2014, 43% of HHs nationally were cell only (Blumberg & Luke, 2014)	Depends on coverage of existing survey	Depends on coverage of existing survey
Response rate	Low, less likely to get "buy-in" for survey over the phone	May be high, although the presence of interviewers in certain locations may act as deterrent. If interviewers are perceived as law enforcement, there may be cooperation issues.	Moderate
Need for visuals	Can be provided to respondents only if mailing precedes RDD call	High degree of flexibility in use of visuals	Can be provided with mail materials, or in advance of web survey invitation
Size of geographic area	No impact	Depends on range of existing survey but large geographic area may be difficult to cover and/or prohibitively expensive	No impact
Screening costs	High, may need to screen HHs to identify fishers or people who consume fish	Low - screening may be unnecessary, or minimal	Moderate, would have to screen to identify fishers or people who consume fish, but less costly than phone or in-person screening

DRAFT

Table 3-5. Level 1 framework for choosing a sample design and data collection mode based on independent factors (continued)

	RDD Telephone Survey	In-Person Intercept Survey	Web or Mail Survey
Literacy rates among the target population	No effect on ability to participate in survey	No effect on ability to participate in survey	Those with low literacy may be unable to participate, leading to potential biases
Web penetration for target population	No effect on ability to participate in survey	No effect on ability to participate in survey	No effect on mail survey, major impact on web survey
Telephone coverage for target population	Persons with no landline numbers would be unable to be sampled or contacted	No effect on ability to participate in survey	No effect on ability to participate in survey
Questionnaire length (see Section 4.9 for details)	Screener up to 5 min., 20-30 min. for main survey, longer lengths can be used but potential for breakoffs will increase	10-20 minutes, but situational factors need to be considered	Web 15-20 mins., but higher potential for breakoffs with longer length; Mail 15-30 mins, although up to 45 min. may be feasible
Relative costs	Low - Moderate	Moderate - High	Low

* This refers to how many have an email address (proxy for web access) or have a telephone number available.

DRAFT

Table 3-6. Level 2 framework for choosing a sample design and data collection mode based on independent factors

YES = APPROPRIATE MAYBE = MAY BE APPROPRIATE NO = NOT APPROPRIATE	LEVEL 2: LOW TO MODERATE COST OPTIONS FOR DEVELOPING A NEW SURVEY						
	MAIL ONLY		MAIL TO WEB		WEB ONLY		IN-PERSON
	Address Based-Sample	List Sample	Address-Based sample	List Sample	Address-Based sample	List Sample	Intercept Survey
SURVEY OBJECTIVES							
What response rates are needed from the survey to meet precision requirements?							
High (≥80%)	NO	MAYBE	NO	MAYBE	NO	MAYBE	MAYBE
Moderate (60-79%)	MAYBE	MAYBE	MAYBE	MAYBE	MAYBE	MAYBE	MAYBE
Low (<60%)	YES	YES	YES	YES	YES	YES	YES
Are visuals needed for validity (e.g., river maps, pictures of fish species, portion sizes, etc.)?							
Yes	MAYBE	MAYBE	MAYBE	MAYBE	YES	YES	YES
No	YES	YES	YES	YES	YES	YES	YES
Are data needed over multiple seasons or are multiple observation points needed?							
Yes	YES	YES	YES	YES	YES	YES	MAYBE
No	YES	YES	YES	YES	YES	YES	YES
What is the size of the geographic area to be covered?							
Large	YES	YES	YES	YES	YES	YES	NO
Moderate	YES	YES	YES	YES	YES	YES	MAYBE
Small	YES	YES	YES	YES	YES	YES	YES
POPULATION CHARACTERISTICS							
Prevalence of characteristic or behavior of interest among target population (e.g., pregnant women and/or high fish consumers)?							
High percent of population	YES	YES	YES	YES	YES	YES	YES
Moderate percent of population	MAYBE	MAYBE	MAYBE	MAYBE	MAYBE	MAYBE	MAYBE
Low percent of population	NO	MAYBE	NO	MAYBE	NO	MAYBE	NO
Unknown	MAYBE	MAYBE	MAYBE	MAYBE	MAYBE	MAYBE	MAYBE
What are the literacy rates among the target population?							
High	YES	YES	YES	YES	YES	YES	YES
Moderate	MAYBE	MAYBE	MAYBE	MAYBE	MAYBE	MAYBE	YES
Low	NO	NO	NO	NO	NO	NO	YES
Unknown	NO	NO	NO	NO	NO	NO	YES

DRAFT

Table 3-6. Level 2 framework for choosing a sample design and data collection mode based on independent factors (continued)

	MAIL ONLY		MAIL TO WEB HYBRID		WEB ONLY		IN-PERSON
	Address-Based sample	List Sample	Address-Based sample	List Sample	Address-Based sample	List Sample	Intercept Survey
How important is it to have high coverage of persons with no fixed address?							
Very important	**NO**	**NO**	**NO**	**NO**	**NO**	**NO**	**YES**
Moderately important	**MAYBE**	**MAYBE**	**MAYBE**	**MAYBE**	**MAYBE**	**MAYBE**	**YES**
Not important	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**
What are the web penetration rates for the target population? *							
High	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**
Moderate	**YES**	**YES**	**MAYBE**	**MAYBE**	**MAYBE**	**MAYBE**	**YES**
Low	**YES**	**YES**	**NO**	**NO**	**NO**	**NO**	**YES**
Unknown	**YES**	**YES**	**NO**	**NO**	**NO**	**NO**	**YES**
What are the telephone coverage rates for the target population? *							
High	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**
Moderate	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**
Low	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**
Unknown	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**
PROS/CONS OF EACH APPROACH							
Coverage	Highest coverage of population (excluding area probability samples). Potential for higher rates of undercoverage in some rural areas.	Will not represent a general population. The extent to which a targeted population will be covered will depend upon list quality.	Highest coverage of population (excluding area probability samples). Potential for higher rates of undercoverage in some rural areas. Depends on web penetration.	Will not represent a general population. The extent to which a targeted population will be covered will depend upon list quality. Depends on web penetration.	Will not represent a general population unless web penetration is 100% among population.	Will not represent a general population. The extent to which a targeted population will be covered will depend upon list quality. Also depends on web penetration.	Difficult and costly to design statistically representative survey; excludes individuals who fish from private ponds.

DRAFT

Table 3-6. Level 2 framework for choosing a sample design and data collection mode based on independent factors (continued)

	MAIL ONLY		MAIL TO WEB HYBRID		WEB ONLY		IN-PERSON
	Address-Based sample	List Sample	Address-Based sample	List Sample	Address-Based sample	List Sample	Intercept Survey
Response rate	Moderate		Moderate		Low		May be high, although if interviewers are perceived as law enforcement, there may be problems with obtaining cooperation.
Need for visuals	Can be provided with mail materials		Variable, can be included in mail survey or shown on web survey		Can be shown on web survey		High degree of flexibility in use of visuals
Size of geographic area	No impact		No impact		No impact		Large geographic areas may be difficult to cover and/or prohibitively expensive
Screening costs	Moderate, may have to screen to identify fishers/ people who consume fish, but less costly than phone or in-person screening	Generally not necessary, but will depend on target population in the list sample	Moderate, may have to screen to identify fishers/ people who consume fish, but less costly than phone or in-person screening	Generally not necessary, but will depend on target population in the list sample	Moderate, may have to screen to identify fishers/ people who consume fish, but less costly than phone or in-person screening	Generally not necessary, but will depend on target population in the list sample	Low - screening may be unnecessary, or minimal
Literacy rates among the target population	Those with low literacy may be unable to participate, leading to potential biases		Those with low literacy may be unable to participate, leading to potential biases		Those with low literacy may be unable to participate, leading to potential biases		Little or no impact

DRAFT

Table 3-6. Level 2 framework for choosing a sample design and data collection mode based on independent factors (continued)

	MAIL ONLY		MAIL TO WEB HYBRID		WEB ONLY		IN-PERSON
	Address-Based sample	List Sample	Address-Based sample	List Sample	Address-Based sample	List Sample	Intercept Survey
Web penetration for target population	No effect on ability to participate in survey		Sampled persons without Internet access will be unable to fully participate, leading to potential biases		Sampled persons without Internet access will be unable to participate, leading to potential biases		No effect on ability to participate in survey
Telephone coverage for target population	No effect on ability to participate in survey						
Questionnaire length (refer to Section 4.9 for details)	15-30 mins., but up to 45 min. may be feasible		Mail invitation only, no data collected; Web- 15-20 mins., but higher potential for breakoffs with longer length		15-20 mins., but higher potential for breakoffs with longer length		10-20 minutes, but situational factors need to be considered
Relative costs	Low		Low		Low		Moderate (with relatively small sample size)

* This refers to how many have an email address (proxy for web access) or have a telephone number available.

DRAFT

Table 3-7. Level 3 framework for choosing a sample design and data collection mode based on independent factors

YES = APPROPRIATE MAYBE = MAY BE APPROPRIATE NO = NOT APPROPRIATE	LEVEL 3: MODERATE TO HIGH COST OPTIONS FOR DEVELOPING A NEW SURVEY						
	TELEPHONE		IN-PERSON			MAIL TO PHONE HYBRID	
	RDD	List Sample	Area Probability Sample	Address-Based Sample	List Sample	Address-Based Sample	List Sample
SURVEY OBJECTIVES							
What response rates are needed from the survey to meet precision requirements?							
High (≥80%)	NO	MAYBE	MAYBE	MAYBE	MAYBE	NO	MAYBE
Moderate (60-79%)	MAYBE	MAYBE	MAYBE	MAYBE	MAYBE	MAYBE	MAYBE
Low (<60%)	YES	YES	YES	YES	YES	YES	YES
Are visuals needed for validity (e.g., river maps, pictures of fish species, portion sizes, etc.)?							
Yes	MAYBE	MAYBE	YES	YES	YES	MAYBE	MAYBE
No	YES	YES	YES	YES	YES	YES	YES
Are data needed over multiple seasons or are multiple observation points needed?							
Yes	YES	YES	YES	YES	YES	YES	YES
No	YES	YES	YES	YES	YES	YES	YES
What is the size of the geographic area to be covered?							
Large	YES	YES	YES	YES	NO	YES	YES
Moderate	YES	YES	YES	YES	MAYBE	YES	YES
Small	YES	YES	YES	YES	YES	YES	YES
POPULATION CHARACTERISTICS							
Prevalence of characteristic or behavior of interest among target population (e.g., pregnant women and/or high fish consumers)?							
High percent of population	YES	YES	YES	YES	YES	YES	YES
Moderate percent of population	MAYBE	MAYBE	MAYBE	MAYBE	MAYBE	MAYBE	MAYBE
Low percent of population	NO	MAYBE	NO	NO	NO	NO	MAYBE
Unknown	MAYBE	MAYBE	MAYBE	MAYBE	MAYBE	MAYBE	MAYBE

DRAFT

Table 3-7. Level 3 framework for choosing a sample design and data collection mode based on independent factors (continued)

	TELEPHONE		IN-PERSON			MAIL TO PHONE HYBRID	
	RDD	List Sample	Area Probability Sample	Address-Based sample	List Sample	Address-Based sample	List Sample
What are the literacy rates among the target population?							
High	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**
Moderate	**YES**	**YES**	**YES**	**YES**	**YES**	**MAYBE**	**MAYBE**
Low	**YES**	**YES**	**YES**	**YES**	**YES**	**NO**	**NO**
Unknown	**YES**	**YES**	**YES**	**YES**	**YES**	**NO**	**NO**
How important is it to have high coverage of persons with no fixed address?							
Very important	**NO**	**NO**	**MAYBE**	**NO**	**YES**	**NO**	**NO**
Moderately important	**MAYBE**	**MAYBE**	**MAYBE**	**MAYBE**	**YES**	**MAYBE**	**MAYBE**
Not important	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**
What are the web penetration rates for the target population? *							
High	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**
Moderate	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**
Low	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**
Unknown	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**
What are the telephone coverage rates for the target population? *							
High	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**
Moderate	**MAYBE**	**MAYBE**	**YES**	**YES**	**YES**	**MAYBE**	**MAYBE**
Low	**NO**	**NO**	**YES**	**YES**	**YES**	**NO**	**NO**
Unknown	**MAYBE**	**MAYBE**	**YES**	**YES**	**YES**	**MAYBE**	**MAYBE**
PROS/CONS OF EACH APPROACH							
Coverage	Excludes HHs w/o landline telephone numbers, although could supplement with cell frame.	Will not represent a general population. The extent to which a targeted population will be covered depends upon list quality.	Highest coverage of population.	Highest coverage of population. Potential for higher rates of undercoverage in some rural areas.	Will not represent a general population. The extent to which a targeted population will be covered depends on list quality.	Highest coverage of population. Potential for higher rates of undercoverage in some rural areas.	Will not represent a general population. The extent to which a targeted population will be covered will depend upon list quality.

DRAFT

Table 3-7. Level 3 framework for choosing a sample design and data collection mode based on independent factors (continued)

	TELEPHONE		IN-PERSON			MAIL TO PHONE HYBRID	
	RDD	List Sample	Area Probability Sample	Address-Based Sample	List Sample	Address-Based Sample	List Sample
Response rate	Low, less likely to get "buy-in" for survey over the phone. Potential for lower response for cell phone respondents.	Variable, depending upon list quality and data collection mode.	Relatively high			Relatively low; Potential survey "buy-in" increases w/previous mail contact.	Variable, depending upon list quality and data collection mode.
Need for visuals	Can only be provided through advance mailing to respondents.		High degree of flexibility in use of visuals.			Can be provided only through advance mailing to respondents.	
Size of geographic area	No effect on ability to participate in survey.		Large geographic areas will require sample clustering.		Large geographic areas may be difficult to cover and prohibitively expensive.	No effect on ability to participate in survey.	
Screening costs	High, may have to screen to identify fishers or fish consumers.	Generally not necessary, but will depend on target population in the list sample.	Variable, depends on if only interested in fishers or fish consumers.		Generally not necessary, but will depend on target population in the list sample.	Moderate-High, if screening is part of mail protocol. If screening part of phone protocol, screening costs will increase.	Generally not necessary, but will depend on target population in the list sample.
Literacy rates among the target population	No effect on ability to participate in survey		No effect on ability to participate in survey		Little or no impact	Those with low literacy may be unable to respond to phone number request, leading to potential biases.	

DRAFT

Table 3-7. Level 3 framework for choosing a sample design and data collection mode based on independent factors (continued)

	TELEPHONE		IN-PERSON			MAIL TO PHONE HYBRID	
	RDD	List Sample	Area Probability Sample	Address-Based Sample	List Sample	Address-Based Sample	List Sample
Web penetration for target population	No effect on ability to participate in survey.						
Telephone coverage for target population	High if cell phones included.	Those with no telephone number on the list frame will be unable to be contacted.	No effect on ability to participate in survey.			Sampled persons without telephones would be unable to provide a phone number for contact.	Those with no telephone number on the list frame will be unable to be contacted.
Questionnaire length (see Section 4.9 for details)	20-30 min., although longer lengths can be used but potential for breakoffs will increase.		30-60 minutes		30-60 minutes	Mail invitation only, no data collected; phone 20-30 mins. with higher potential for breakoffs with longer length.	
Relative costs	Moderate		High		High	Moderate	

* This refers to how many have an email address (proxy for web access) or have a telephone number available.

DRAFT

4 QUESTIONNAIRES FOR COLLECTING FISH CONSUMPTION DATA

4.1 Background

We have previously presented a variety of data collection options (based on budgetary level) in Tables 3-5 through 3-7. These tables address various types of sampling frames and modes that can be used, based on available resources and research goals. In addition to selecting a sampling frame and mode to satisfy the goals of the planned research, a variety of factors such as characteristics of the population (e.g., literacy rates) and desired outcomes of the survey (e.g., response rates) also factor into the decision. These are also addressed in Tables 3-5 through 3-7 to provide guidance for the researcher. In this chapter, we build upon this information by presenting questionnaire approaches that can be used with any of the survey approaches previously discussed. This is the fourth of the five basic steps in the design and development of a survey for estimating fish consumption rates listed in Chapter 1– Development of an appropriate survey questionnaire. Note that this chapter does not cover methodologies for determining heritage rates. Those methodologies are covered in Chapter 6. As additional resources, Appendix A provides a discussion of considerations for instrument development to help the researcher develop survey questions that are methodologically sound. Appendix B provides examples of validated survey instruments, as well as additional information on structuring questions aimed at collecting fish consumption data. See also Willet, 2012, and Thompson and Subar, 2008, for general information and discussion on dietary assessment.

The decision of which dietary assessment tool to use depends on many things. Two of the most important considerations are the ultimate objective of the survey (e.g., estimate the distribution of usual intake) and the planned analysis method. The NCI has published the Dietary Assessment Primer, which provides background on dietary data collection methodologies, the benefits and drawbacks of the various methods, and approaches to use given the study objective (https://dietassessmentprimer.cancer.gov/). If the survey's goal is to estimate the usual intake distribution, the NCI's Diet Assessment Primer endorses multiple administrations of a 24-hour recall or a single 24-hour recall on the full sample and multiple 24-hour recalls on a sub-sample (https://dietassessmentprimer.cancer.gov/approach/table.html#intake). NCI notes that more than two administrations of the 24-hour recall are beneficial for foods that are consumed by less than 5 percent to 10 percent of the population of interest. However, further research is needed to guide specific recommendations on the number of 24-hour recalls needed in relation to frequency of consumption. They therefore suggest that the best approach is to collect two 24-hour recalls from the entire sample (https://dietassessmentprimer.cancer.gov/approach/intake.html).

In Section 1.2.3, five analysis methods were introduced to estimate UFCR from dietary data collected through a survey. These methods are the NCI Method (Tooze et al., 2006; Tooze et al., 2010), the Multiple Source Method (MSM) (Harttig et al., 2011), the Iowa State University Method (Nusser et al., 1996), the Statistical Program for Age-Adjusted Dietary Assessment (SPADE) (Dekkers et al., 2014), and the use of a FFQ. The use of an FFQ alone, not in combination with 24-hour recalls, is cautioned. This method results in more biased estimates compared to the NCI Method as FFQs result in more measurement error and are more prone to systematic error than 24-hour recalls, which are used for the NCI Method (Kipnis et al., 2003). Table 4-1 provides the data

needs and the advantages and limitations of each of these analytical approaches. Other analysis methods requiring custom programming might be used, including Bayesian methods. In some cases where the data are not adequate for the NCI method, additional data may be available that can be used to estimate UFCR, in particular see Carroll et al., 2012. In these cases, consultation with a statistician is recommended.

The NCI Method for data analysis utilizes statistical modeling to estimate usual intake of nutrients and foods, and it is currently the best method for estimating usual intake of episodically consumed foods such as fish. The method has a general requirement that dietary data are collected at two or more different time points, during which a respondent provides a 24-hour recall. The method provides an estimate of UFCR representing the long-term average grams of fish consumed per unit of time (e.g., day or week). There is a requirement that at least some respondents have reports of consumption of the fish of interest for more than one time point. It is generally accepted that a survey must yield at least 50 respondents who report fish consumption at least two different times (Kipnis, et al., 2009). In order to achieve this requirement, a survey must cast a wide net, especially if the prevalence of the behavior of interest (e.g., consumption of a certain type of fish) is low. While lengthening the recall period at each interview time point per contact (e.g., 24 hours to 7 days) increases the number of fish consumption reports, the accuracy of a participant's memory of what he or she ate diminishes as the recall period lengthens (Gersovitz et al., 1978).

If there is a concern that the survey will not yield enough reports of fish consumption from individuals on multiple days for certain subpopulations or specific fish of interest (e.g., freshwater fish) to successfully implement the NCI Method, it would be beneficial to also collect long term frequency of consumption data from participants. It may be possible to approximate the distribution of fish consumption in small subpopulations or for rarely consumed fish species based on an analysis of the relationship among mean fish consumption rates from the FFQ and the distributions of fish consumption in those populations where it can be estimated. As a simple example, if a survey yields enough reports of multiple days of total fish consumption to utilize the NCI Method, but not enough reports of multiple days of freshwater fish consumption, the UFCR distribution for freshwater fish might be estimated by scaling the UFCR distribution for all fish based on the ratio of the mean fish consumption rate from the FFQ data.

Table 4-1. Data needs, advantages, and limitations of various analytical methodologies for estimating UFCR

	NCI Method	Multiple Source Method (MSM)	Iowa State University Method	Statistical Program for Age-Adjusted Dietary Assessment (SPADE)	FFQ Alone
Data needs	24-hour fish consumption recall	24-hour fish consumption recall	24-hour fish consumption recall	24-hour fish consumption recall	Frequency of consumption over a long period of time; some measure of portion size
Minimum number of contacts with respondents	2	2	2	2	1
Advantages	Current best method for estimating usual intake of episodically consumed foods like fish. Can utilize covariates to improve estimates (such as frequency of consumption). Can include survey weights.	Can utilize covariates to improve estimates (such as frequency of consumption). Web-based interface is easy to use.	Can include survey weights.	Implemented in the statistical package R and freely available. Can include age as a covariate. Can include survey weights.	Least expensive due to dietary data collection requirements and easiest to implement. Can include survey weights.
Limitations	Requires SAS software (expensive)	Does not provide standard errors. Cannot include survey weights.	Does not account for correlation between consumption amounts and consumption frequency.	Does not account for correlation between consumption amounts and consumption frequency.	Estimates of UFCR may be significantly biased. The amount of bias is hard to quantify and will be different for estimates of upper and lower percentiles.

4.2 Questionnaires for Fish Consumption Surveys

Defensible and quantifiable fish consumption rates are needed for development of water quality standards for tribes and states, as well as assessment of seafood contaminant risks to the population. There are various options that can be considered when choosing a questionnaire for a fish consumption survey. In the overview provided in Appendix B, the methods of a 24-hour recall, fish focused recall, food record/diary, food frequency questionnaire, and fish diet screener are described. Depending on the needs of the survey, any of these may be appropriate for use. Appendix B enumerates the advantages and limitations of each approach.

For purposes of this discussion, we focus on a previously used instrument that utilizes the 24-hour recall approach and also collects fish consumption frequency data. In 2003, EPA developed a Microsoft Access-based software package, henceforth called the "Fish Consumption Survey Tool" (or FCST), which could be used to conduct well-designed fish consumption surveys (Kissinger et al., 2010). While the software, derived from a 2000 Suquamish Tribe seafood consumption survey and other seafood consumption surveys used in the Pacific Northwest, was originally developed for tribal use, it is also usable by other entities and other populations.

The FCST is sophisticated and well developed and provides reasonably good estimates of fish consumption. Multiple administrations allow for the assessment of usual intake and population distributions (based on the NCI analysis method described in Chapter 7). It also requires the respondent to estimate the portion of fish consumed, as would be the case in all retrospective surveys.

The FCST is a configurable, computerized interview that collects a 24-hour fish consumption recall, annual and seasonal consumption (including source and parts consumed), purchased versus caught fish, consumption by children, and consumption at gatherings and special events. Its greatest strength is the capture of quantitative fish consumption data. It allows researchers to populate the software with fish species, parts, portion sizes (including dishes) of interest and to characterize seasonal variability in fish consumption. Fish consumption may be characterized on the basis of groups of fish having similar feeding behavior/trophic level associated with contaminant uptake. The fish database includes 77 species of fish. The FCST automates logic flow in question branching, simplifying survey administration. Range checking of answers enhances accuracy and reliability. The need for data entry of hard copy surveys is eliminated. Individual interviewers can use the FCST on multiple devices and can download their data to a master survey database. Query and data reporting capabilities allow calculation of basic seafood consumption statistics on either a complete survey data set or data subsets. The data could also be used with the NCI Method for estimating UFCR. The survey includes a booklet with images of a variety of species, portion amounts and fish preparation methods. The questions that collect more qualitative information have been well vetted, although users may wish to modify these questions. A drawback of the FCST is that modification of these questions may require the services of a Microsoft Access programmer and associated consultants.

As it stands, the automated version of the FCST can be conducted in-person or via telephone using a desktop computer, laptop, or other mobile device. If an automated survey is not possible, then the interview can also be implemented as a hard-copy questionnaire. This would allow for the mail data collection mode to be used. Additionally, using the hard-copy questionnaire as a guide, the

instrument could be converted into a web survey. Links to both the automated version and a hard-copy version are available on EPA's website (https://www.epa.gov/fish-tech/epa-guidance-developing-fish-advisories), along with a User's Guide, Supervisor's Guide, and other accompanying documentation (copies of the guides, documentation, and hard-copy questionnaire are provided in Appendix D). These resources are being provided to researchers free of charge by EPA. Other survey instrument options are presented in Appendix B.

4.3 Use of Visual Aids

When consumption of fish from specific geographic areas is to be assessed, maps of the area of interest can be provided to ensure that respondents understand the questions. Use of a map helps to standardize the frame of reference when reporting local products. Geographic information systems resources are readily available at the federal, tribal, state, and county levels.

As local names for fish may vary, it may also be helpful to provide photographs of the fish species of interest, keeping in mind that some consumers may not be familiar with the appearance of the intact fish. Inquiries should be performed with local experts prior to finalizing the wording of the survey to make sure that common local names are used when possible.

Visual aids are commonly used to help respondents estimate portion sizes when collecting dietary intake. These aids may include life size photographs of relevant portions of various fish preparations, physical portion size models, or common objects. Lately, there has been interest by dietary researchers in using common objects as portion aids, such as a checkbook or deck of cards, rather than the traditional measuring cups and spoons. Appendix D contains information on preparing photographs on portion size.

It should also be noted that misidentification of types of fish can present a major obstacle for consumption surveys that are interested in a limited set of species. To reduce misidentification, visual aids are often used to assist the respondent with correct identification of fish consumed. This step is a challenge for telephone surveys and other interviews conducted without the presence of an interviewer. If the interview is not conducted in-person, procedures should be implemented for sending, or otherwise making available, visual aids such as maps and fish species photographs for use during the interview.

4.4 Respondent Body Weight

The body weight of survey respondents is an important variable in analyzing fish consumption data. Tribal, state, and local governments often have an interest in reporting fish consumption on a per-body weight unit basis, as lighter individuals may be able to consume less fish before they reach safety thresholds for contaminants in fish. In order to understand variability in fish consumption within a population, it can be useful to collect and then control for the weight of the survey respondents. If study respondents are, on average, heavier or lighter than the overall population, then the consumption rate may differ (Trondsen et al., 2004). Such differences can only be

accounted for only if respondent weight is collected. If respondents are expected to report on their entire household, then the weights of all household members in question should also be captured.

Although it is known that self-reported weight can be inaccurate (Connor et al., 2007), obtaining self-reported weight is still a generally accepted survey practice. Using data from the National Health and Nutrition Examination Survey (NHANES), it has been found that men tend to overreport, and women tend to underreport, their weight. Variations can also be seen depending on the age of the respondent and whether or not he/she is at a normal weight, underweight, or obese (Merrill & Richardson, 2009).

CDC's BRFSS survey includes a question on self-reported weight: "About how much do you weigh without shoes?" Average weights for populations within states are available from the BRFSS annual data bases (http://www.cdc.gov/brfss/annual_data/2013/pdf/CODEBOOK13_LLCP.pdf). These weights could be used if the population of interest is the state-wide general population.

4.5 Portion Sizes

The amount of fish consumed is the main variable of interest when estimating UFCR. It is important to understand that the terms "portion" and "meal" are not necessarily equivalent. For example, a meal may consist of shrimp consumed as an appetizer and baked bass as a main course. When collecting the portion size data from the respondent, the survey team needs to differentiate between the fish species in this example and collect a portion size for the shrimp and a separate portion size for the bass. Additionally, if the same species is consumed more than one way (i.e., fried shrimp and boiled shrimp) then the respondent needs to provide portion size data for both preparations.

Generally, respondents estimate the size or amount of the fish that they consumed (for example, the size of a deck of cards or a checkbook or a half a cup). For mixed fish dishes, the weight of fish per unit volume may be ascertained from recipes and combined with estimates of consumption volume to yield estimates of fish intake. These amounts are converted to gram or ounce equivalents for analysis.

In order for respondents to more accurately estimate their portions, visual cues are used. Pictures of fish portions of varying sizes (e.g., 4 oz., 6 oz.) on standard size plates help respondents to estimate the amount of fish that they consumed. As discussed in Section 4.3, if the interview is not conducted in-person, these pictures of fish meal sizes may be sent (or otherwise made available) to respondents for their use during the interview.

While it is ideal for respondents to provide the amount they consumed in raw weight, they may not know the weight of the fish unless they were the one who bought or prepared it. Many surveys collect the as-consumed weight or amount and convert it to raw weight during analysis (see Section 7.3.1). It is important for the survey to consistently collect either raw or as-consumed amount or to ask and record which the respondent is providing.

If respondents are providing portion sizes for as-consumed fish, it is necessary to ask how the fish was prepared (e.g., breaded, battered) and how it was cooked (e.g., poached, fried, baked). This will

allow for conversion from as-consumed amounts to raw weight if that is the estimate of interest. See Section 2.3.4 for a list of preparation attributes that may be of interest.

If the survey is using the FCST, and a respondent does not know the weight of the fish he/she consumed, it is designed so that a respondent views a photograph of a fish portion (e.g., a baked salmon filet) that has a known raw weight (e.g., 4 oz.) and then estimates the weight of the fish they consumed. This requires the survey team to weigh and cook fish meals of species of interest to photograph for use in the survey. See Appendix C, Developing the Booklet of Photographs, for more information.

4.6 Collecting Data by Fish Species

A survey that needs to collect data for multiple species of fish needs to be aware that asking respondents to estimate frequency of consumption by species can lead to overestimation of fish consumption. For example, individuals may overestimate their salmon consumption if they are asked about consumption of many individual salmon species. This overestimation is likely lower in cultures where fish consumption is important and members of the population are very familiar with consumption of specific salmon species. An alternative approach could be to ask about frequency of consumption of a group, e.g., salmon, and then ask the respondents to provide a percentage breakdown of the individual species they consume (Chinook, Coho, pink, sockeye, etc.).

Another consideration regarding collecting data by species is that some surveys may be interested in a factor that varies by species. For example, a survey may be interested in knowing what fraction of the fish that is consumed is harvested. This is quite difficult if some species of fish are almost always harvested and others are almost always purchased. The respondent has to engage in a difficult cognitive exercise to come up with assignment of fish source percentages. Thus these types of questions are better posed on a species specific basis, for example, "what fraction of salmon are harvested or caught by you or someone you know?"

4.7 Fishing Practices and Locations

The variety of fishing practices, whether it is recreational catch and release, recreational catch and consumption, or subsistence-based, will affect a respondent's report of consumption. It will be important to collect fishing behavior preferences from those respondents who report fishing behavior. This can include whether they consume the fish they catch, preferences for type(s) of fish they may fish for and subsequently consume or give away, and locations respondents may prefer.

For on-the-bank or intercept surveys it will be important to evaluate locations frequented by fishers. Some locations may be avoided by fishers due to an advisory or information from other fishers that some locations may be undesirable. If fishing locations are not well identified and missed, this may result in some groups or minority populations not being covered by the survey, especially if they tend to avoid well-populated areas. For location-based surveys, good sources for identifying locations are park officials or personnel from state departments of fish and game. It is also

important to note that if the objective of the study is to estimate unsuppressed FCR, then conducting surveys at locations with fish advisories would not be preferable.

Advisories, creel limits, and fisher observation of caught fish will influence fishing practices and subsequent consumption rates (for example, see Beehler et al., 2001). Collecting information on these behaviors can help to identify possible reasons for consumption suppression and may allow for quantification of suppression for specific populations of fishers.

4.8 Data Capture and Management Methods

A number of interrelated factors can help to determine the optimal data capture method for a consumption survey, of which the most important may be data collection mode, questionnaire complexity, sample size, and available resources to process data.

For self-administered mail questionnaires, a paper-and-pencil instrument can be used to collect the data from respondents; once the questionnaire is returned the answers need to be captured in a way that facilitates analysis. If the number of questionnaires is small, then it may suffice to have a staffer extract the questionnaire data. However, as the number of questionnaires grows and/or the complexity of the questionnaire increases, so, too, does the effort necessary to capture the data and, in turn, the likelihood of coders introducing error into the data capture process. A computerized data-capture system can help to reduce the time necessary to code a large number of questionnaires while reducing the likelihood of introducing error into the process through coding and key entry errors. A computerized data-capture system requires sophisticated software that can accurately read respondents' answers on each page. The FCST offers a computerized format.

For self-administered web-based questionnaires, data capture is automatic as the respondent inputs their responses into the computer directly. Up-front time and costs are required to upload and/or program the questionnaire; however, once this is complete, the survey can be replicated across respondents and survey administrations (e.g., seasons) and results can be extracted at any point. If a self-administered web-based survey is planned, web accessibility by the target population should be taken into account.

Questionnaire complexity should be a consideration when using self-administered (i.e., paper and pencil) instruments as there is no interviewer to assist respondents when they're confused or if they make a mistake. Particularly with self-administered mail questionnaires, respondents must be able to understand what the questions are asking of them, which questions apply to them, and how to appropriately answer and navigate the instrument. If this is not clear to the respondent, then the questionnaire may yield inaccurate and/or incomplete data, or in some cases result in item or respondent nonresponse. An advantage of a web-based or computerized instrument is that the questions respondents see or hear can be tailored based on their responses to previous items. For example, respondents who indicate they have not consumed fish can be directly moved past likely followup questions on fish consumption. Such tailoring is not possible in a paper-and-pencil questionnaire but can be approximated by including clear instructions to respondents about which questions they should answer based on their previous answers. However, unlike a web or computerized survey, respondents cannot be forced to adhere to such instructions and data cleaning will be necessary, along with procedures to address errors of commission or omission. When

transcribing data from paper forms into electronic data storage, it is important to implement checks on transcription quality.

For surveys conducted over the phone or in-person, it is the interviewer who must capture the data provided by respondents. Similar guidelines as described above for self-administered questionnaires apply to interviewers capturing data; it may suffice to have the interviewer fill out a paper-and-pencil questionnaire if the questionnaire is short and simple and the resources are available to extract the data and enter it into an analyzable format. Similar to a web-based questionnaire, a Computer-Assisted Telephone Interviewing (CATI) program can be used to facilitate more complex questionnaire administration and simple and error-free data processing.

As mentioned, the FCST (discussed in Section 4.2) can be used as a computerized instrument, although a hard-copy version is also available. Based on the availability of computers and programming resources that may be available for a particular study, a decision will need to be made as to whether automation is feasible. Computerized data collection (vs. capturing data on hard copy) is usually desirable and results in higher quality data, due to the ability to automatically conduct edit checks during the data collection process and directly input data into a file.

4.9 Interview/Questionnaire Length

It is widely accepted that interview/questionnaire length is correlated with a respondent's likelihood to participate in the survey. The longer the survey, the greater the burden and the less likely respondents will be willing to participate. Groves et al., 1992, cites interview length as a deciding factor in the decision to participate in face-to-face surveys. However, Bogen, 1996, has found the influence of interview length on participation behavior to be "weak and inconsistent." Instead, the effect of interview length on survey participation may actually be related to timing of the survey request or the perception of length. In a Danish study, Hansen, 2007, observed increases in participation when the expected length told to the respondent was shortened from 20 minutes to 15 minutes (which reflected an actual decrease in survey length) for a telephone interview. The question for interview length then becomes, how long is too long? Other factors should be considered, including the relevance, or saliency, of the survey topic. Tolerance levels of participants for lengthier interviews may be increased if the subject matter is particularly salient or important to them.

Another concern with interview length is the effect that a long or burdensome interview can have on data quality. Holbrook et al., 2003, found that as respondents got to the end of a long telephone interview, they observed more satisficing behaviors. Respondent satisficing can take many forms, but generally involves respondents taking cognitive shortcuts to select the first acceptable (or easiest) response. Satisficing behaviors can lead to substantial measurement biases (for more on satisficing, see Krosnick, 1991).

While it can be tempting to researchers to develop an exhaustive questionnaire covering a wide range of topics of interest, it is critical that the questions included in the survey instrument map to research goals or objectives. Researchers need to consider the response burden placed upon the participant to ensure that quality data are collected. It is also important to consider the population that will be interviewed when thinking about interview length. For example, creel or on-the-bank intercept surveys may need to be shorter as fishers may be unwilling to provide time for a lengthy

interview after a day of fishing (Washington State Department of Ecology, 2012). Conversely, the narrative structure of sharing or conveying information in many Native American cultures may lead to much longer interviews and for these situations guided or structured interviews may be more appropriate as opposed to strict adherence to written questions.

4.10 Qualitative Methods to Support Survey Instrument Development

Qualitative research is beneficial during the questionnaire development process or development of other survey components (e.g., location maps, identification of fishing locations, fish pictures). This will help ensure that survey questions appropriately map to the research objectives. Testing will also help identify difficult questions and concepts or ambiguous terms that are not familiar to the respondent. Key informants that are familiar with local or customary fishing practices, culture, and fish preparations can be identified and assist at this stage. Terminology can be clarified through this qualitative research, to ensure that data being collected are those which were intended. Names for certain species of fish, or how to identify fish that are bought versus those that are caught, are just a few of the types of clarifications that can be ascertained. There are commercial software packages available that are useful for analyzing qualitative data.

In using qualitative research methods, we must carefully consider what is necessary to reach the research goal. These methods can, for instance, be used to identify access points and locations (for a specific body of water) that may be frequently used. Fishers may use specific points for early fishing hours but may migrate to a different location as water temperature changes. To understand this type of pattern and ultimately build a frame for a creel or on-the-bank survey, the researcher could perform cognitive interviews on a broad spectrum of fishers. Frequent fishers' behavior may differ from that of recreational fishers. The researcher should also provide materials so the participants can identify locations (on a map, for example).

When the objective is to define how to measure consumption, qualitative research methods can be useful to help identify how people identify various types of fish or how they think of consumption. Frequent fishers may identify fish more easily by sight or be familiar with the types of fish they catch. Recreational fishers may better recognize specific names for fish, or local names, rather than other names/labels that are used.

A difficult question to answer is how many respondents are needed for qualitative research approaches. This type of research can be done in groups or on an individual basis. Identifying fishing locations or how people identify fish is an activity that can be done in a group or individual setting. Testing questionnaires is something that should be done in a one-on-one setting. The number of participants can vary but should not be so many that it overwhelms the research team. The number also should be sufficient enough to identify common themes or problems that can be distinguished from idiosyncrasies that may not be an indicator of a real problem. For more guidance on qualitative testing methods, see Willis, 2005.

4.11 Pilot Testing

After the survey instrument and accompanying procedures have been developed, pilot tests are conducted to identify possible shortcomings within a survey instrument or procedures that generally cannot be identified during qualitative testing or expert review. A pilot test is a small-scale test of survey instruments and procedures conducted under actual survey conditions (e.g., within a realistic environment using respondents similar to the survey's targeted respondents). It may also serve to identify unique interviewing issues specific to the populations of interest. For example, a common interviewing technique when a respondent does not understand the question is to repeat the question. This may present a cultural issue in some populations, for example when a younger tribal member is interviewing a tribal elder, as the younger tribal member may feel that they are insulting the intelligence of an elder. Ideally, a pilot test can be done early enough in the survey period to allow for small changes, as needed, to the questionnaire and/or study procedures prior to the start of actual data collection. Data resulting from pilot testing are typically not included as part of the analytic data set, and the individual respondents interviewed during the pilot test are typically not interviewed during the actual study.

Pilot tests can vary in size and scale, depending on resources and time available, but they should always have set goals or objectives before they are conducted. For example, a pilot test can be conducted to test new procedures, or for data collection efforts involving multiple processes (e.g., multiple contacts, diaries, etc.), to ensure that all appropriate procedures are in place and operating as planned. A pilot test can also be used to help estimate sample sizes (by obtaining information about response) or prevalence of a behavior in the population. Finally, pilot testing can be useful to help identify to what degree respondents are able to provide information desired by the researcher. For example, if a researcher is interested in knowing the consumption of a specific type of fish, pilot testing can help identify whether respondents are aware of the fish species they consume.

Ideally, the pilot test will try out not only the questionnaire but also give a thorough workout to production or field procedures, including methods of finding participants and use of data capture instruments (paper and pencil, computer-assisted telephone interview , or computer-assisted personal interview), as well as coding and key entry procedures. The pilot test of the questionnaire should include pilot respondents who will give all aspects of the questionnaire a thorough workout. The following list of characteristics is an example of participants who might be recruited for a pilot test of a questionnaire in an ethnic minority group.

- Age: elders and younger participants
- Gender: male and female
- Fishing: fishers and non-fishers
- Source of fish: primarily eat at home vs. eat out frequently
- Income: low-income and medium- or high-income;
- Food preparation: persons who do and who do not usually prepare food for themselves or the household
- Recent immigrants vs. members of the ethnic group who have lived in the U.S. for longer periods of time
- Members of all of the various ethnic groups that might be encountered in the survey

It is also important to include individuals across the spectrum of lifestyle from traditional to modern.

A given pilot participant may satisfy more than one of the criteria noted above. For example, someone may be an elder who follows a traditional lifestyle and fishes. Generally, a minimum of 10 persons should be recruited as pilot participants, although more may be needed. Note that for Federally-funded research, approval by the Office of Management and Budget (OMB) may be required and if so, pilot testing is strongly encouraged, but may be limited to 9 or fewer respondents thus not requiring formal approval of an information collection request (ICR). https://www.whitehouse.gov/sites/default/files/omb/inforeg/statpolicy/standards_stat_surveys.pdf

Sequential improvement of the questionnaire after every participant or after several participants during the pilot is a process that ends when no more changes to the questionnaire or procedures are needed. It is likely that some minor changes to the questionnaire or procedures will be needed during the main survey, but if the pilot is well done, these changes should be minimal.

5 SURVEY IMPLEMENTATION AND OPERATIONAL CONSIDERATIONS

5.1 Introduction

Once the sampling frame, sample size, mode of administration, and type of survey instrument have been established, there are important considerations for fielding a study. The quality of a survey is dependent on the quality of the data collected. Data collection is usually the most labor intensive aspect of a survey and also the most expensive. Therefore, it is important to consider implementation and operational issues during the development process. This is the fifth and final of the basic steps in the design and development of a survey for estimating fish consumption rates listed in Chapter 1 – Consideration of implementation and operational issues. This chapter describes some of the key issues. There are many other factors to consider in operationalizing a survey, including hiring and training of interviewers, interviewer supervision, interviewer assignments, etc. These topics are not within the scope of this document; however, it is important for the researcher to consult with a survey specialist to address these issues.

5.2 Defining Eligible Respondents

For operational purposes, it is necessary to determine who is eligible to respond to the survey. Once the target population has been determined, consideration should be given to situations in which sampled persons might not respond for themselves. Below, several common situations are discussed.

5.2.1 Household Reporters

If the survey will be collecting information about all eligible persons in the household (this includes both screening surveys and consumption surveys that ask about the behaviors of each eligible household member), it is important for study procedures to specify who is eligible to serve as a respondent for the household. (See Section 3.3 for a discussion of within-household sampling, in which a sampled person is selected from among all eligible persons in the household.) Generally, surveys require that household respondents be adult members of the household. More restrictive requirements might necessitate it be the primary respondent for the household (e.g., the person most knowledgeable about the household members' consumption).

5.2.2 Children

If the target population includes children, it will be necessary to determine who is eligible to respond for the child. Young children themselves may not be able to serve as respondents due to both data quality concerns and the need to obtain parental consent. Researchers should also be advised that there are special regulatory requirements generally accepted across Federal, state, and local agencies

that provide additional protection for children who would be involved in research. Additional information can be found at http://www.hhs.gov/ohrp/policy/populations/children.html.

However, with carefully considered procedures, food consumption surveys do sometimes obtain information from children directly, depending on their age. For example, NHANES collects food intake data for children under the age of 6 from a proxy respondent; proxy-assisted interviews are conducted with children 6-11 years of age; and participants 12 years and older complete the dietary interview on their own.

In many cases, the respondent for a child is specified as the parent or guardian of the child. For fish consumption surveys, it may be preferable to require that the respondent be a parent or guardian of the child who is knowledgeable of the child's food intake. In shared custody arrangements or other situations in which the child resides in more than one location, other special considerations may come into play when collecting consumption information.

5.2.3 Proxy Respondents

There may be other situations in which proxy respondents are necessary. For example, if a person is disabled or elderly and cannot respond for themselves, a person who is knowledgeable about their intake may be interviewed. While proxy respondents can be used to collect information about food consumption, the concern is with data quality. However, in some situations, the only alternative to proxy reporting might be nonresponse. When weighing whether to permit proxy respondents, the researcher should consider what particular situations might give rise to a need for proxy respondents (i.e., a preference for proxy respondents over complete nonresponse). Consideration should also be given to what stipulations might be placed on the potential proxy respondent (e.g., a requirement that he/she is knowledgeable of the sampled person's food intake).

5.3 Pre-contact Procedures and Introduction of the Study

Careful pre-contact procedures, along with a persuasive introduction to the survey, can serve to increase study participation. The use of pre-notification letters is a useful method for introducing the study and alerting potential respondents of future contact by the survey organization. Research has shown that pre-notification letters can lead to increased response in mail surveys (Dillman et al., 2009), RDD telephone surveys (de Leeuw et al., 2007; Link & Mokdad, 2005), and telephone surveys utilizing lists (Goldstein & Jennings, 2002). For in-person surveys, these letters may be delivered by the interviewer to provide the respondent additional information or amplify study legitimacy.

The content of the letter should include an appeal for participation, highlight the study sponsor, convey the importance of the study, and provide information for what is expected of the respondent. Letters that are lengthy (generally more than one page) or provide an overabundance of information should be avoided, as respondents are unlikely to read letters that appear burdensome. For more information on letter content and design, see Dillman et al., 2009.

The timing of the delivery of the letter is important so that it directly precedes the substantive encounter (e.g., a mailed paper survey or interviewer contact). If the letter is mailed too far in advance, the connection between the pre-notification letter and the survey contact may be forgotten. Pre-notification letters should be sent a few days to a week in advance of the survey or initial interviewer contact.

Additionally, eliciting support from community leaders for surveys of targeted groups can also be important for the success of a survey. Including endorsements of the survey by community leaders in communications may enhance survey response and credibility. Identification of the primary media utilized by the target population and preparation of communications designed for those media to publicize the survey may also be helpful. See, for example, a fish consumption survey for Asians and Pacific Islanders in King County (U.S. EPA, 1999).

5.4 Informed Consent for Interview

Informed consent is the process by which the prospective research participant is provided with sufficient opportunity and information about the study to decide whether or not to participate. Ethical standards specify that researchers may not conduct research on human subjects unless they have obtained consent from the respondent or the respondent's proxy. Depending on the type of data collection, oral or implied consent may be appropriate. For in-person interviews, written consent is needed.

For Federally-funded studies, as well as many state and local surveys, planned consent procedures must be reviewed and approved by an Institutional Review Board (IRB). IRB review is also important and sometimes required for non-Federally funded studies. The IRB will ensure the protection of human subjects based on regulations from the Office for Human Research Protections (OHRP) within the U.S. Department of Health & Human Services. If the project will include protected populations such as pregnant women or children, additional review may be required, and review boards will require assurance that the study procedures include protection of such groups. Additional information about regulatory requirements that govern research on human subjects may be found at http://www.hhs.gov/ohrp/. If appropriate for a specific research endeavor, some consideration should be given to using IRBs that have experience with the unique cultural, social, and psychological impacts on Native Americans.

The IRB of record for the study will require information on the informed consent procedures to be followed. For many study designs such as mail or telephone data collection, some informed consent procedures will be impractical. The IRB can waive requirements for written informed consent procedures, but generally only under the following conditions:

- Research poses minimal risk to study participants;
- Subjects' rights or welfare will not be affected by participation;
- It would be otherwise impractical to conduct the research; and
- Full disclosure of the risk and benefits of participation will be provided to participants.

Although written informed consent may be waived in certain instances, respondents must always be informed of their rights as research participants. If approved by the IRB, verbal consent may be

used to inform them of these rights. Informed consent requires that the consent to participate in research is informed, understood, and voluntary. Informed consent must be conducted in a manner that minimizes coercion or undue influence. These ethical principles are in the Belmont Report and have been codified in the Federal regulations to protect human research subjects, found at 45 C.F.R. § 46.116. Researchers must ensure that informed consent documents meet the regulatory requirements, including state and local laws where the study takes place. The following nine required elements of informed consent must be communicated to a potential research subject.

1. Identify the project as research, explain its purpose, and state the expected duration of the subject's participation
2. Describe the procedures to be followed
3. Describe any foreseeable risks (i.e., probability of physical, psychological, social, or economic harm or injury occurring as a result of participating in the study)
4. Describe any potential benefits to either the individual or society that can be expected from participating in the research
5. Disclose alternative medical treatments, if any
6. Explain how the confidentiality of any data that identifies the subject will be maintained
7. Provide the names and contact information for two people—one to contact to ask questions about the research and another to contact about the subject's rights as a research participant
8. State that participation in the research is voluntary
9. Explain whether compensation or medical treatment will be available if any injury occurs (required for more than minimal risk studies)

IRBs may vary in having additional requirements, and there may be additional criteria for protected populations. Researchers are encouraged to meet with their IRB early in the project design to establish necessary requirements for approval. Fish consumption surveys may often be timed to coincide with specific seasons; therefore, it is important that adequate time is allocated for appropriate review of the project before any data collection activities are scheduled to begin.

Tribes are sovereign nations and researchers are required to obtain specific permission from the (or each) Tribe to conduct research. In addition to the study IRB, the study should be prepared to apply for a tribal research permit, which may require seeking approval from a separate tribal IRB or the tribal government. The Indian Health Service (IHS) has regional IRBs, such as the Portland Area Indian Health Board. Some tribes use other IRBs such as the Northwest Indian College's IRB. This will vary from tribe to tribe. In many cases, studies will need to seek an agreement to the use of the intellectual property generated by the study from the tribe and be prepared to share all information collected with the tribal authority. In addition to IRB requirements, tribes may also have provisions that require the research activity does not damage their cultural heritage or identity.

Additionally, EPA has information on human subjects review that can be accessed at https://www.epa.gov/osa.

For additional information, the reader is encouraged to review a thorough discussion on this topic by Harding et al., 2012.

5.5 Impact of Seasonality on Data Collection Schedule

As mentioned previously, the timing of the data collection will affect consumption rate estimates. Ideally, a survey would capture seasonal variation in fish consumption by planning on a full year of data collection. However, this may not be feasible. Questions determining how soon information from the study will be needed, or over what seasons the data are needed, should be considered to help determine when the study should be conducted. Fishing activity might be undertaken by the majority of fishers only during the summer; however, ice fishing is popular in some areas of the northern United States. In addition, fish caught in one season might be preserved (e.g., smoked or frozen) and consumed later, indicating that exposure to tissue contaminants might be equally important year-round. If the data collection period misses the peak seasonal periods, an underestimate of consumption would result. Conversely, measuring consumption only during seasonal peaks may result in an overestimate of consumption. One way to address seasonality is to divide the sample into equal proportions and interview each over fixed time periods until the entire sample has been interviewed. Unequal numbers of interviews per time period can be accommodated by weighting approaches. Another approach might be to conduct interviews during two periods, one of low consumption and one of high consumption. The research goals should dictate when and, possibly, how often consumption data will be needed.

5.6 Language and Literacy Issues

Low literacy rates can be problematic if prevalent within the target population or an area that has been sampled. The effectiveness of some study designs can be impacted by low literacy rates. For example, self-administered survey designs (mail, web) will likely see a negative impact on survey response. This can result in biased estimates if fish consumption behavior is correlated with literacy.

The National Center for Education Statistics (NCES) provides state and county estimates of low literacy, defined as the percent lacking Basic Prose skills in the National Assessment of Adult Literacy (NAAL), which includes those who score Below Basic in prose and those who could not be tested due to language barriers (NCES, 2003). NCES has an online tool that provides state and county estimates of the percentage of the population lacking basic prose literacy skills (http://nces.ed.gov/naal/estimates/stateestimates.aspx). In some counties in the United States, the illiteracy rate is over 30 percent (NCES, 2003).

Literacy of the population should be considered both when developing survey questions and in choosing a mode of administration. If a self-administered survey mode is under consideration, estimated literacy rates should be reviewed for the target population and taken into consideration at the study design phase. For study designs where interviewers administer the survey, there may be self-administered components (e.g., food log or diary) that could be affected by literacy rates. Interviewers should be trained to identify respondents who may have difficulty with these components. In many cases, respondents may be reluctant to reveal literacy issues to the interviewer. In the design of any survey questions (as well as consent forms, letters, etc.), the designer should strive to keep language to less than a 12th grade level. See Fowler, 1995, and Willis, 2005, for further guidelines on question design.

If the target population or the sampled area has a high percentage of non-English speakers, it is important to initially review the prevalence of other languages spoken by members of the target population. The language differences may be addressed through the use of questionnaires translated into the predominant non-English language(s) spoken for self-administered surveys. For interviewer-mediated surveys, it will be important to have trained data collectors who speak the predominant languages within the sample. Best practices in survey translation should be followed in order to tailor questions to the needs of a given audience while retaining the measurement properties of the source. Planning for translation should be part of the study design (Harkness et al., 2010).

5.7 Survey Nonresponse

Survey response is a critical component of any survey and will have potentially strong influence on measurement bias, sample size, and costs in terms of data collection effort. Trends for response rates have shown an overall decrease in response over time for household surveys (Brick & Williams, 2013). This same effect has been observed for natural resource based surveys (Connelly et al., 2003). It is highly recommended that an experienced survey statistician be involved in planning for nonresponse analysis, as it is not possible to proscribe a "one-size-fits-all" approach within this guidance document. The application of methods for analyzing and ameliorating the effects of nonresponse are extensive and require a solid background in survey methods as well as training in nonresponse analysis.

Nonresponse is unavoidable and should be anticipated. Plans for dealing with nonresponse should be a part of the study design and analysis methods. To minimize nonresponse, approaches for encouraging cooperation should be built into the survey design. These may include multiple contact attempts with selected respondents, as well as follow-up reminders for nonresponders such as postcards, letters, and telephone calls, if feasible. For self-administered surveys, generally mail surveys, replacement surveys should be sent to nonresponders (Dillman et al., 2009). Following protocols outlined by Dillman and his colleagues, replacement surveys are mailed to sample units that have not responded to an initial survey. For interviewer-mediated surveys (e.g., in-person or telephone surveys), this will require procedures for re-contacting the household and specialized training for interviewers attempting to convert nonresponders.

The survey mode will have an effect on the response rate that can be anticipated. In-person surveys generally have the highest response but also the highest per interview costs. Mail surveys often outperform or do as well as telephone surveys in terms of response rates. However, a telephone survey may be necessary for complicated interviews. Web surveys generally have the lowest overall response and often lack an available sample frame, but web surveys mixed with mail invitations to complete the survey have shown promise (Millar & Dillman, 2011).

The degree to which nonresponse has a biasing effect will depend upon the relationship between the propensity to respond to the survey and the characteristic of interest from the survey (Groves, 2006). If fish consumers have a higher propensity to respond than do non-consumers, this will result in higher per capita consumption estimates. (Per capita consumption estimates include both consumers and non-consumers as part of the population.) This will not be an issue if consumers only are the population of interest (i.e., if non-consumers are not part of the target population for the survey). However, if nonrespondents have a higher consumption rate than respondents, this will

result in an underestimation of fish consumption in both per capita and consumer-only consumption rates.

Also, it should be noted that avidity bias can be a problem for fish consumption surveys. This can arise when sample members disproportionately participate due to interest, perceived relevance, or activity in the survey topic. For example, recreational fishers may be more likely to participate, but family members or others who also consume the sport-caught fish may feel the survey is of less interest or relevance.

Determining the degree to which nonresponse bias may be present requires planning for a nonresponse bias study. While nonresponse analyses may not always be feasible for smaller or lower budget surveys, it may be of interest to note that for studies requiring approval by the Office of Management and Budget (OMB), standards dictate that a nonresponse bias analysis be done when unit or item nonresponse suggest the presence of bias, e.g., if response rates fall below 80 percent (OMB, 2006). Any nonresponse analysis should be done in consultation with a survey expert.

The purpose of the nonresponse study is to detect whether there is a difference in the behavior of interest between those who responded to the survey and those who did not. The number of cases necessary will be driven by the estimate of the behavior for respondents. If the behavior of interest is of relatively low frequency, it may be necessary to sample and interview a larger pool of nonrespondents in order to have enough power to detect a difference between the two groups. For studies with small sample sizes, a nonresponse bias study may not be feasible.

Generally, nonresponse bias studies include a subset of the survey items that are related to the measure of interest or that will generate estimates. In large studies where the pool of nonrespondents is large, cases from the pool of nonrespondents are subsampled, since these cases may have previously received multiple contact attempts and will require much additional effort and costs to contact all nonrespondents. Where possible, it is best to attempt a different mode of contact. For example, if the survey was conducted by telephone, an in-person contact with nonresponders may be more successful.

5.8 Methods for Respondent Retention

For some study designs it may be necessary to re-contact survey participants to complete data collection. In the context of collecting data on fish consumption, an example of a design where re-contact will be necessary is when multiple dietary recalls are completed (e.g., separate 24-hour or multiple-day recalls). It will be necessary to have procedures in place for re-contacting respondents to collect the second recall.

In order to mitigate respondent attrition for study designs where re-contact will occur over a short period of time, researchers should inform the respondent of plans for future contact. In the case of reluctant respondents, it can be helpful to inform them of the purpose of the later contact. Additionally, it is beneficial for retention purposes to set a time or an appointment for the next contact during the first data collection, although this could affect respondents' consumption behavior for the day they know they will be reporting about. An appointed day/time for the interview can help commit the respondent to the task and identify optimal times for re-contact.

Finally, if possible, it may be helpful to have the same interviewer conduct the first and subsequent data collections, as research has shown this can help with respondent retention (Hill & Willis, 2001).

For study designs that will follow respondents over a longer period of time (e.g., multiple seasons or years) to measure individual-level change in consumption over seasons, for example, one or more intermediate contacts with the respondent may be necessary. Intermediate contact (i.e., contact between substantive data collections) will help to keep the respondent involved in the study, be reminded of past cooperation, and help maintain up-to-date contact information. At baseline, it is important to obtain alternate contact information for each respondent, including alternate telephone numbers, e-mail addresses, and/or contact information for close relatives or friends. This information is useful in the event that original contact information becomes outdated during the data collection period. A protocol for interim contacts between data collection events should be developed based on the study schedule and budget, and can include one or more of the following approaches: periodic telephone or e-mail reminders to the respondent, reminder postcards regarding upcoming activities, and/or letters that could include information about the study progress, as appropriate. To ensure adequate sample sizes for analysis, it is very important to retain respondents in the study who have previously participated, and careful attention to retention efforts should take place as part of overall study planning..

5.9 Confidentiality

Maintaining the confidentiality of information or responses provided by respondents is important to the integrity of all surveys. Respondents will be asked to voluntarily provide data that they may generally be unwilling to share with others. While fish consumption can appear to be an innocuous behavior or topic, under some circumstances it can be quite sensitive. For example, pregnant women may be advised to avoid certain species of fish or shellfish. Respondents may not want others to be aware of how much (or how little) fish they consume. Assuring respondents that their information and privacy is of utmost importance to the researcher can help foster cooperation with the survey request.

Many state agencies offer legal protections for the confidentiality of survey data and also may have specific requirements for researchers to follow. Researchers should verify they are in compliance with any state, local, or tribal regulations for maintaining respondent confidentiality. Additional protections can be granted for surveys in the form of certificates of confidentiality that are available through several government agencies (for example, see https://grants.nih.gov/grants/policy/coc/index.htm).

Assurances of confidentiality provided to respondents are an important part of gaining cooperation and trust during the interview process. All survey information must be kept confidential, and any plans to share data must be explained to participants during the informed consent process prior to data collection. Steps must be taken to separate survey data from personal identifying information (PII), and procedures for doing this must be described in the IRB submission(s). PII should be removed from the data files. At a minimum, this involves replacing identifiers, such as names, with anonymous ID numbers. For public use data files, more intensive procedures involve data disclosure analysis to verify respondents cannot be identified due to a combination of sampling (e.g., geographic area) and their unique survey data. Procedures for anonymizing data are available that

preserve the estimates generated by the survey data (see https://fcsm.sites.usa.gov/committees/cdac/ for more information).

Each person working on the study must be continuously aware of his or her responsibility to safeguard survey data and identifying information. Names or other information about study participants should never be divulged to anyone except the research team. Data collectors should be instructed to never interview someone that they know personally. This may lead to response bias and a possible breach of confidentiality. All members of the research team are under the same legal, moral, and ethical obligations to protect the privacy of those participating in the research. All members of the research team, including data collectors, should sign a Pledge of Confidentiality that is prepared by the survey team. Additionally, specific training on confidentiality and privacy procedures may be required for all staff working on a research project. For more information related to confidentiality and tribal communities see Harding et al., 2012.

6 SUPPRESSION

As described in Section 2.3.5, a suppression effect occurs when a fish consumption rate (FCR) for a given population, group, or tribe reflects a current level of consumption that is diminished from an appropriate baseline level of consumption for that tribe, population, or group. For most survey goals and objectives, it is important to assess suppression, in particular if the resulting FCR will be used in regulatory action. Additionally, if one objective of the survey is to investigate consumption trends over time, and there are no survey data available from comparable populations, the techniques presented in this chapter can be helpful in estimating past consumption.

There are a variety of methods and approaches that can be used to estimate an unsuppressed FCR. One approach is to recreate the historic, ecological conditions through the use of historical records and documents and apply knowledge of nutrition, natural resource use, and activity levels to estimate the historic FCR or heritage rate. In many cases, heritage rates may be the only practical way to estimate unsuppressed rates – that is, free from the biasing influence of suppression effects. While this approach could be particularly useful for tribal fishing populations, it may not be feasible for other populations or situations where there is a lack of necessary historical information or other data. Alternative methodologies for establishing unsuppressed rates are presented in Section 6.2.

6.1 Estimating Heritage FCR

Heritage rates of fish consumption can be validly estimated by detailing the ecological conditions under which the population lived and how the population utilized the available natural resources. The subsistent (or traditional, historic) diet can then be recreated and the fraction of that diet which was fish can be determined. (Harper & Walker, 2015a, 2015b; Harper et al., 2012; Harper et al., 2007)

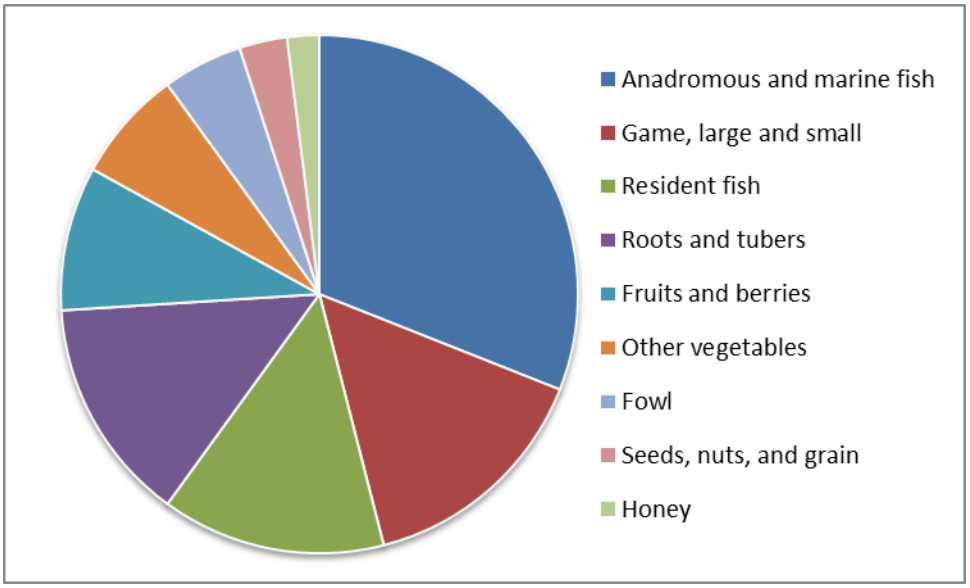
The data and information can come from a variety of sources. The use of historic, ecological data can provide the types and amounts of natural resources that were available to the population during the time period of interest and how they varied throughout the year. This information includes ecoregion maps, climatological data, watershed data, and species composition data. Ecoregion maps and detailed descriptions of the regions can be found at http://www.epa.gov/wed/pages/ecoregions.htm. The National Oceanic and Atmospheric Administration National Climate Data Center provides access to historical climate data (available at: https://www.ncdc.noaa.gov/cdo-web/) and paleoclimatology data (available at: http://www.ncdc.noaa.gov/data-access/paleoclimatology-data/datasets). The United States Geological Survey has watershed maps that are available at http://water.usgs.gov/wsc/map_index.html. Species information can be found in state Natural Heritage Programs (for example see http://www.azgfd.gov/w_c/edits/species_concern.shtml) and archeological records can be used to determine historical populations.

Anthropological and ethnographic documents can provide the type and amount of activities undertaken by the population, how the natural resources were used, and the cultural significance of these natural resources. Some populations may have historical written accounts of their fishing activities, including quantities harvested and consumed, in which case this may be enough

information to estimate the heritage FCR. The population's traditional environmental knowledge (TEK) identified through cultural interviews, ethno-historical and ethno-botanical literature, oral history, and historical records, such as those from early observers, can also provide the basis for the determination of historical or traditional resource use patterns. A great deal of valuable information can be gained through discussing experiences and oral histories with tribal members, especially tribal elders.

From the ecological data and the anthropological and ethnographic information identified above, a food pyramid or food wheel can be constructed that represents the subsistent diet and shows the relative importance of the identified dietary staples, such as fish, game, roots, honey, etc. An example is shown in Figure 6-1 (modified from Harper et al., 2007).

Figure 6-1. Food wheel showing the relative proportion of dietary staples in an example subsistent diet



Note that one population may have more than one subsistent diet depending on the ecological findings. For example, three subsistent diets have been identified for the Wabanaki – inland freshwater, inland anadromous, and coastal based on three major ecological settings in the state (Harper & Ranco, 2009; Harper et al., 2007). Other examples are available at http://superfund.oregonstate.edu/conducting-research-tribal-communities.

A nutritionally complete diet can be constructed that shows the grams per day of each food consumed by determining the representative foods in each staple category (e.g., for anadromous and marine fish it may be salmon, oysters, and herring and for game it may be deer, rabbit, and beaver) and assigning caloric counts to each (Kcal/g) from the USDA National Nutrient Database for Standard Reference (USDA, 2010). For example, if it was determined that anadromous and marine fish comprised 30 percent of the total diet, and the average Kcal/g given the species mix was 1.55 Kcal/g, we can estimate that to reach a daily calorie goal of 2,000 Kcal, a typical individual would have consumed 600 Kcal (30% of 2,000 Kcal) of anadromous and marine fish each day, which is 387 g/day (600 Kcal divided by 1.55 Kcal per g). It is important for researchers to clearly document

each calculation and estimate at each step, providing the assumptions and documenting decisions made.

To obtain information on the uncertainty of an estimated heritage rate, the researcher can vary the assumptions and develop rates under each set of assumptions. While this range is not as formal as a confidence interval, it will give the user some feeling for the uncertainty in the rates.

Unlike the exacting and fully specified methodology for a contemporary survey of current consumption, the heritage rate estimates may depend on professional judgment. That judgment may be informed not only by the literature, but also by site visits to the population and use of historic ecological data. Heritage rates should not be discarded simply because professional judgment or a reasonable assumption has been used in the calculations. The user of the rates take into consideration the underlying basis of the heritage rate.

The methods described in this section provide a peer-reviewed approach (Harper et al., 2012) for estimating an unsuppressed FCR. Additionally, examples of studies include the following (posted at http://superfund.oregonstate.edu/conducting-research-tribal-communities):

- Confederated Tribes of the Umatilla Indian Reservation (CTUIR) (Harris & Harper, 2004)
- The Spokane Tribe's Multipathway Subsistence Exposure Scenario and Screening Level RME (Harper et al., 2002)
- Washoe Tribe Human Health Risk Assessment Exposure Scenario for the Leviathan Mine Superfund Site (Harper, 2005)
- Quapaw Traditional Lifeways Scenario (Harper, 2008)
- Wabanaki Traditional Cultural Lifeways Exposure Scenario (Harper & Ranco, 2009)

Although this methodology provides a reasonable estimate of an unsuppressed rate, it cannot be assumed that this approach could be reproduced for all populations. This depends on the types of historical records and/or data that are available for the population.

6.2 Alternative Methodologies to Measure Suppression

In some cases, a similar population may be divided among two areas, one contaminated and one not. To the extent that one area can be a control site for another area, those two populations might be compared, assuming that the populations are truly similar. In other cases, a single population might have had data collected before a significant contamination event, in which case the population serves as its own before- and after- control. These types of studies are typical epidemiological study designs, and are described in environmental epidemiology texts along with ways to control for confounding factors.

If they are available, fish consumption rate data from areas with similar fish and human populations to the area of interest that are less impacted by fish advisories or fears of chemical contamination can be used. This method measures suppression by comparing one contemporary site to another control site. This is not the same as suppression in fish consumption from a traditional or heritage

rate. The control site is a baseline with respect to the general population, but is not a baseline for a tribal heritage rate. A key factor in this decision is the habitat quality associated with areas from which the population harvests fish. This approach has been employed in examining chemical contaminant risks for American Indians/Alaska Natives consuming seafood in the vicinity of Puget Sound Superfund sites (U.S. EPA, 2007). The suppression effect due to the fish advisories or the fears of chemical contamination could then be estimated by taking the difference between the two rates. Note that the resulting estimate of the suppression effect would not account for any suppression in the less impacted area (unrelated to fish advisories or chemical contamination), such as suppression from a heritage rate.

A similar methodology to assess suppression due to changing fish availability could be investigated by looking at the relationship between fish populations and fish consumption rates. For example, in areas with plentiful fish, what are the FCRs for the population and how do those FCRs compare to FCRs in areas with fewer fish? Investigating these differences could provide some insight into the magnitude of suppression due to changing fish availability, but will not measure any suppression from the heritage rate.

Another appropriate alternative is to include suppression questions in a fish consumption survey. For example, a person might want to eat more fish, but they are not readily available and/or expensive or a family might want to eat more fish if they had time to fish. These questions can be designed to determine whether fish consumption may be lower than a population would normally consume or wish to consume. Suppression questions can take several forms, but should be developed in such a way so that the causes and reasons behind suppression can be identified. Since the measure of suppression relies upon recall of past (in terms of years or decades), or even historical, consumption patterns, studies often incorporate more qualitative or ethnographic approaches.

Quantification of suppression can be very difficult using traditional survey methods, but could be addressed at a qualitative level through use of well-structured questions. One principle in the design of survey questions is to beware of hypothetical questions as people generally are unprepared to predict future behavior or quantify past behavior, especially if it has been a generation or more since suppression began. Estimating future behavior is difficult as there are often a number of influences or situational factors that a hypothetical question is unable to account for. Another reason that responses to hypothetical questions can be troublesome is that the question or questions are often based on hypothetical changes in a state or behavior that the respondent must create or imagine. These imagined states may be difficult for a question to fully describe. (Fowler, 1995).

From a methodological perspective, it is important that good standard survey practices be employed if these topics are important for a proposed survey. In particular, respondents may be asked, "If the fish were not contaminated, how much would you eat?" or "How much fish would you like to eat if you could do so safely?" These types of questions are speculative and based on hypothetical situations only. Estimates derived from such speculative questions may yield data that say little or nothing about actual future behavior or about heritage cultural practices.

However, if a survey requires that questions about future behaviors be asked, it is generally appropriate that the questions be based as much as possible on past experience or direct knowledge by the respondent. This helps to provide some frame of reference for the respondent. For example, if respondents are asked if they plan to continue current fish consumption levels in the future,

current fish consumption habits and experiences should be included in the questioning. This is in contrast to simply asking respondents if their consumption habits would change based on hypothetical changes in other factors (e.g., availability or safety). The respondents may have no past experience on which to base their response. When this occurs, respondents are likely to make cognitive shortcuts in the response process, rely on heuristics, or provide socially desirable answers.

Some approaches to asking about current behaviors/experiences as they may relate to future activities include the following:

- Determine if the respondents have tried to eat fish but the fish they wanted was unavailable
- Ask about the awareness and influence of fish consumption advisories on fish consumption
- Ask about economic influences that may prevent them from eating fish (e.g., cost of purchased fish relative to income)
- Ascertain the importance of fish as a dietary staple or the cultural significance of fish consumption
- Ask about the current availability of the fish within the respondents' local area
- Ask whether a family or respondent's access to fishing locations has changed
- Ask whether laws or fishing regulations have altered the ability of the respondent to obtain fish
- Ask whether changes in family duties related to working in the contemporary job market have reduced the time family members have to engage in fishing activities

While responses to these questions will not be able to quantify a hypothetical change in a behavior, they can indicate whether a change in any of these states or situations may be a factor in influencing future behavior.

Another approach to capture suppressed or time-altered consumption is to ask respondents whether their fish consumption or fishing activities have increased, decreased, or stayed the same in comparison to a period X years ago (e.g., 10 years). Note that long-term suppression or contamination means that the heritage, unsuppressed, or baseline rate cannot be ascertained by asking these types of questions because they may not remember or have not studied the issue of population-specific traditional rates. Such a question—with a fixed prior time point—cannot be asked of the younger respondents. Only respondents who were teenagers or older at the prior time point can give meaningful answers. For respondents who indicate a change in consumption, the reason for the change can be captured. The FCST (the EPA developed Microsoft Access-based software package) includes the question, "*Sometimes for various reasons, people's consumption of fish and shellfish changes. Has the amount of fish and/or shellfish consumed by you or your family changed over the last twenty years?"* and the follow-up questions, "*How has the amount of fish and/or shellfish you have eaten over the past 20 years changed? (eat more now, eat less now, eat different types now)* and *Please tell me what you think has caused the change in the amount or type of fish and shellfish you now eat*." These questions can be used to determine if there is a suppression effect, gauge the possible magnitude of it, and find ways to adjust the FCR that results from the survey for the suppression effect. For example, if respondents report that they are consuming less fish now and the reasons provided are related to environmental contamination, comparing the resulting FCR from the survey to a FCR from an uncontaminated area could provide an adjustment factor. Another example would be if respondents reported that they eat less fish now

due to the availability of fish in lakes and streams for harvest, researching changes in fish populations could provide an adjustment factor.

It is recognized that, in some cases, it may be necessary or possibly unavoidable to ask hypothetical questions. When these situations arise, careful development of such questions is encouraged. The area of survey methodology called "contingent behavior" is based on the concept that people are capable of accurately predicting their future actions, given well-worded questions about specific scenarios (Parsons et al., 2006).

Although quantitative data about future consumption often cannot be derived from the questionnaires described in this guidance document, by asking people to predict their level of future use under the change of a single condition (e.g., alleviation of their concerns about contamination), a survey can provide useful information on the qualitative scale of change that usage rates are likely to undergo as remediation and/or risk communication progresses.

7 ANALYTIC APPROACHES

7.1 Introduction to Analysis

This section provides a general discussion on statistical analysis and Section 7.2 focuses on the types of data analysis available for estimating UFCR using the FCST described in Chapter 4, or other similar dietary data collection tools (e.g., providing both 24-hour fish consumption recall data and frequency of fish consumption data).

7.1.1 Weighting

Care must be taken when analyzing the data because the statistical methods appropriate for calculating unbiased estimates of the population parameters will depend on the sampling method (e.g., simple random sampling, stratified sampling, proportional stratified sampling). There are important statistical issues to consider when making adjustments for the various types of sampling inaccuracies. Weights might need to be applied during the estimation of population parameters whereby the weights account for different sizes of subpopulations, for differential nonresponse rates, or for disproportionate sample selection probabilities. For example, there might be cases where probabilities of respondent selection become disproportionate in field implementation such that the sample population disproportionately represents different demographic groups. In those cases where probabilities change between the design and implementation stages, post-stratified weights are used to estimate population parameters that are derived from a sample distribution that does not correspond to the known population distribution. An experienced survey statistician should be consulted to facilitate the appropriate weighted analysis and presentation of survey results.

7.1.2 Estimating Usual Fish Consumption Rates (UFCR)

Estimating the UFCR requires particular dietary data collection and analysis methodologies.

As of the writing of this document, the NCI Method is the most generally accepted method for estimating usual intake. Implementing the NCI Method requires statistical knowledge and the use of SAS software. Individuals experienced with using the NCI Method can be found in academia and statistical consulting and survey research firms. Of the four methods mentioned in Section 1.2.3 for estimating usual intakes using 24-hour recall data (the NCI Method, MSM, the Iowa State University Method, and SPADE), the NCI Method is generally accepted as the method that produces the least biased estimates. Souverein et al., 2011, provide a comparison of the methods. Also see Table 4-1.

The NCI Method (Tooze et al., 2006; Tooze et al., 2010) utilizes statistical modeling to estimate usual intake of nutrients and foods, and it is especially appropriate for episodically consumed foods like fish. It can be used to estimate the distribution of usual intake for a population or subpopulation. The premise of the NCI method is that usual intake is equal to the probability of consumption on a given day times the average amount consumed on a "consumption day." For episodically consumed foods, such as fish, a two-part model is used to estimate usual intake. The

first part uses logistic regression to estimate the probability of consumption with a person-specific random effect. The second part estimates the consumption day amount using linear regression, also with a person-specific effect. The reported consumption amounts are transformed to be roughly normally distributed for linear regression. The regression predictions are then back-transformed to estimate UFCR in the original units (see Tooze et al., 2006 for more details). The two parts are linked by allowing the person-specific effects to be correlated and by including common predictors in both parts of the model. The method assumes all subjects included in the analysis are consumers, even if some are assigned (through modeling) a very small probability of consumption. Thus it produces per capita estimates. Since the method provides long-term averages of consumption, even subjects who report consuming fish once in a year should be included in the analysis. However, if a survey is interested in consumer only estimates, and wants to exclude those who never consume fish (e.g., strict vegetarians and vegans) the survey should include an appropriate question to be able to exclude those from the analysis.

Data from one or more non-consecutive recalls provide the values for the dependent variable. At least a subset of the population, generally believed to be 50 respondents, needs to have consumption data from two or more recalls. Predictors related to either the probability of consumption or consumption amount, such as gender, age, race, and income, can be included in the modeling. In most cases, the most important predictor is a measure of frequency of consumption of the food of interest obtained from a food frequency questionnaire. In addition, the model can incorporate the following within-person predictors: 1) differences between weekends (Friday to Sunday) and weekdays (Monday to Thursday), and 2) consistent differences between multiple 24-hour recalls (sequence effect) as captured by an appropriate covariate. The resulting model parameters are then used to estimate population and subpopulation distributions. Note that research has shown differences in weekend consumption compared to weekday consumption and that Friday is more comparable to Saturday and Sunday than to the rest of the week (Haines et al., 2003). Evidence for the validity of the method has been published in a series of papers in the *Journal of the American Dietetic Association* and *The Journal of Nutrition.* (Freedman et al., 2010; Kipnis et al., 2009; Tooze et al., 2006). The NCI method can be implemented using two SAS programs (the MIXTRAN and DISTRIB macros) available from the NCI website (http://appliedresearch.cancer.gov/diet/usualintakes/macros_single.html).

While the MSM approximates the NCI Method and thus may not provide as accurate estimates of usual fish consumption and does not provide standard errors of estimates, it is available interactively through this website: https://msm.dife.de/, where users can interactively import data sets, define analysis models, and review and export results and graphs. The use of the program is supported by online help and a user guide, and communication between users and the program website is encrypted, securing transmitted data against unauthorized use (Harttig et al., 2011). Using this method may reduce analysis costs and thus may be a good option for some users.

Utilizing an FFQ to estimate UFCR involves asking respondents how often they consume fish in a given time period (i.e., past year, past 6 months) and asking how much they usually consume when they consume it. This amount is generally captured categorically, for example <2 oz., 2 to 6 oz., or greater than 6 oz. (see the Diet and Health Questionnaire available at http://riskfactor.cancer.gov/DHQ/webquest/index.html for an example). To estimate the usual intake for each respondent, the amount reported is multiplied by the frequency reported and then converted to a per day amount. The mean and percentiles of usual intake can then be estimated from these individual rates. Note that this method results in more biased estimates compared to the

NCI Method as FFQs result in more measurement error and are more prone to systematic error than 24-hour recalls, which are used for the NCI Method (Kipnis et al., 2003). It is not recommended for estimating the distribution of usual intake by NCI (https://dietassessmentprimer.cancer.gov/approach/table.html).

In Chapter 4 we discussed dietary assessment methods to collect data for the purpose of estimating UFCR. This chapter describes how the data collected by the FCST should be organized, along with the ancillary data, in order to use the NCI Method for analysis. It also walks through the application of both the MIXTRAN and DISTRIB macros.

7.1.3 Reporting of Usual Fish Consumption Rates (UFCR)

FCR can be reported as "as-consumed" weight or "raw" weight of the fish. There needs to be consistency in how the data are collected in terms of cooked weight (or as-consumed) or raw weight. Depending on the goals of the study, if cooked weights are reported and collected they may need to be converted to raw weights or vice versa prior to analysis. If the as-consumed weight is collected and the data of interest are actually raw weight, it is necessary to collect data on how the fish was prepared and cooked and if the weight is for the whole fish or the edible portion only. Different preparation and cooking methods reduce the moisture in the fish by different amounts. Generally, dry heat (e.g., baking) reduces the weight of the fish by 25 percent, moist heat (e.g., steaming) reduces the weight by 21 percent, frying reduces the weight by 12 percent (U.S. EPA. 2002; U.S. EPA, 2014). The USDA Food and Nutrient Database for Dietary Studies (FNDDS) (USDA, 2014) can be used to determine average recipes for various foods (e.g., breaded and baked trout) to then estimate the weight of the cooked fish in the reported as-consumed amount. If raw weight is the measure of interest, this estimate would then need to be converted to raw weight using a moisture loss conversion factor.

If the survey objectives included estimating UFCR for various sub-populations, then more than one rate will need to be estimated, for example age-specific rates, rates by race/ethnicity, rates by report of sportfishing, etc. If a survey covered a population with a wide-range of different lifestyles, it may be important to estimate and present data by these subpopulations to ensure that subpopulation rates are not masked.

Consumption data can be presented in several different ways. Consumption estimates can be given as point estimates or as distributions illustrating the variability in the population. A point estimate is a single value such as 50 g/day, whereas a distribution can be summarized by a measure of central tendency (e.g., mean, median), a standard deviation, and a shape of the distribution curve (e.g., lognormal). For many risk assessments, risk estimates for individuals at both the central tendency and high-end portions of the exposure distribution are made. The Ambient Water Quality Criteria (AWQC) Human Health Methodology suggests that the arithmetic mean or 90th and 95th percentiles of fish consumption distributions be used for developing AWQC. In characterizing reasonable maximum exposure for Comprehensive Environmental Response, Compensation, and Liability Act (CERCLA) sites, commonly known as Superfund sites, EPA guidance specifies that media intake (e.g., fish consumption) rates be set at the 90th to 95th percentile of the media intake distribution. Estimates should be presented with a measure of precision (i.e. a standard error or 95

percent confidence interval). To preserve the maximum amount of flexibility for future uses of the data, researchers may present consumption data as a distribution.

7.2 Uncertainty

The estimated FCR resulting from the survey and analysis may be uncertain due to either bias (systematic error) or random variation. Bias results in a consistently high or consistently low fish consumption rate relative to the true or desired value. Variation results in an uncertain fish consumption rate that might be either higher or lower than the true value.

The primary sources of random variation are the following:

- Sampling error associated with the random selection of respondents. For example, if different counties and individuals had been selected for the data collection, the data and FCRs would be different
- Random differences due to the simulation of usual fish consumption for each respondent. This source of variation can be reduced by increasing the number of simulations

The confidence intervals for the fish consumption rates account for both of these sources of variation. If there are fewer respondents with reported fish consumption in two (or more) 24-hour recalls, there is less data to estimate the parameters and particularly the variance components, resulting in more uncertainty in the fish consumption estimates and wider confidence intervals.

There are multiple sources of bias that can affect the fish consumption rates including:

- Seasonality
- Respondent bias
- Use of standard recipes to calculate fish consumption amounts from the 24-hour recalls
- Bias associated with the estimation method (e.g., NCI Method, MSM) and its assumptions

Each of these sources of bias is discussed in more detail below.

7.2.1 Seasonality

Fish consumption, especially of recreationally or sport-caught fish, is likely to vary by season. If the data collection plan does not take into consideration seasonality, the estimates may overestimate or underestimate usual intake. Refer to Section 5.5 for more information.

7.2.2 Bias in the Reported Fish Consumption

The reported fish consumption may be biased if respondents tend to report consistently more or less fish consumption in the 24-hour recall than actually occurred. Assessing if the reported values

are biased requires comparing reported values to estimates obtained using other data collection approaches, such as analysis of duplicate meals. Over the years, much research has gone into assessing dietary intake. The OPEN study found that both 24-hour recalls and FFQs yield biased estimates, however FFQs are more prone to systematic error compared to 24-hour recalls and both are prone to random error. Random error can be reduced by greater sample sizes, while systematic error cannot be.

7.2.3 Use of Standard Recipes

The FNDDS can be used to obtain the amount of fish in a recipe (see Section 7.1.3). It utilizes standard recipes for foods such as, breaded and fried trout. If the survey team decides to use this database as opposed to having respondents supply specific recipes of the foods they consumed, the resulting amounts of fish consumed will be biased. For example, the standard recipe for the food "Scallops and noodles with cheese sauce" is approximately 35 percent fish. However, the true recipe for the food consumed by a respondent may have less fish or more fish than the standard recipe. Additionally, there is uncertainty associated with the moisture loss values for processing and cooking methods that will be applied to the amounts to determine raw weight. They are generally average values of moisture loss given the various processing and cooking methods. If respondents cooked their fish a bit longer than the moisture loss would be a bit greater than average, and if they cooked it a bit less, the moisture loss would be a bit less than average.

7.2.4 Estimation of Usual Fish Consumption

Measurements of usual fish consumption are very difficult to obtain. Since usual fish consumption is a long-term average, we would need many 24-hour recalls over a long time to approximate what "usual intake" is trying to assess; therefore we rely on a statistical model and associated assumptions to estimate usual intake. As a result, the estimates of usual fish consumption depend in part on the statistical assumptions.

The statistical model makes certain assumptions, such as, 24-hour recalls provide unbiased estimates of fish consumption, all respondents are fish consumers (at least occasionally, unless they are omitted from the analysis), and the distribution of fish consumption among those reporting consumption in a 24-hour recall is normally distributed for some power transformation. The validity of these assumptions can be discussed and to some extent evaluated using data.

The estimates of the frequency of fish consumption depend in part on how non-consumers (those who never eat fish or don't eat fish for a long time) are treated. From two 24-hour recalls it is not possible to separate true non-consumers from those who did not happen to report fish consumption in either recall. A similar problem relates to consumption of small amounts of fish. Should a person who never eats an identifiable piece of fish but uses a salad dressing with a small amount of fish in it be considered a regular consumer of a very small amount or a non-consumer of fish? Whether a meal is classified as having fish may depend on the procedures used to ask the questions and the recipes used to estimate fish consumption. Having non-consumers in the data will lower the overall probability of fish consumption (P) but increase the variance of the probability of fish consumption

among individuals. The resulting effect on the upper percentiles of the distribution is not clear. The survey team needs to determine if and how they will identify non-consumers to be removed from analysis if desired (e.g., an additional survey question).

The reported amount of fish consumption will vary from one 24-hour recall to another, in part because the respondents may be poor at estimating the amount consumed and in part because the consumption amounts are reported in rounded units, such as a cup or a pint, but not 1.267 cups. The rounding adds some uncertainty to the estimates. The within-person variance component of the statistical models accounts for uncertainty due to poor estimation by the respondent and rounding that is part of the process. Because the definition of usual fish consumption does not include the within-person variation, this source of error should contribute minimal bias to the estimates of usual fish consumption.

The statistical models make some assumptions to simplify the computations, such as an assumption that variance components are normally distributed, additive in the transformed scale, and linearly correlated. The assumption that the person-specific random effect in the probability model is normally distributed is difficult to test without many more 24-hour recalls for each person. The assumption that the two variance components in the amount model are normally distributed is generally consistent with the observation that the Box-Cox transformed consumption amounts are roughly normally distributed. Nevertheless, other assumptions may imply a similar distribution for the reported amounts while using a somewhat different assumption for the person-specific variance component and thus somewhat different estimates of fish consumption. Because the estimated parameters must be consistent with the reported data, the general center and spread of the predicted distribution will be similar regardless of the distributional assumptions. Specific percentiles may be either higher or lower using different assumptions or may be relatively insensitive to the distributional assumptions. Although these assumptions are common in other statistical applications, it is difficult to assess how the estimates might change using other assumptions.

If the model assumptions are accepted as reasonable, then the question is whether the estimates from the model are biased. If the estimates are based on maximum likelihood (as in the NCI Method), this can produce biased estimates, particularly variance estimates, with small sample sizes. Convergence theory says maximum likelihood is best with large sample sizes. Thus, if the sample size is large enough, we expect the estimates to have relatively little bias compared to the size of the confidence intervals.

The fish consumption estimates depend in part on the independent predictors used in the model. When different predictors are used, the estimates change. It is impossible to know what the best set of predictors is. A systematic approach can be used to selecting the independent predictors from the available predictors in an effort to minimize any bias. The estimates have unknown bias due to the decisions that are made.

7.3 Analyzing Recall Data

7.3.1 Preparing the Data

For purposes of this discussion, we will assume that a survey collected two 24-hour fish recalls at least 3 weeks apart and that approximately 25 percent of the participants were contacted during the spring, 25 percent during the summer, 25 percent during the fall, and 25 percent during the winter. Additionally, we will assume that the survey is collecting data to estimate UFCR of fish from rivers and lakes within a specific state, although there may be sub-objectives requiring data to be collected on more than one species.

The fish consumption data collected through the use of the 24-hour fish recall need to be processed and coded such that there are analysis variables consisting of the amount of fish of interest (in our example, fish from rivers and lakes within the state) consumed for each day. Table 7-1 presents example data for five participants. (Note that this is not an example of a data collection form; it is only being used to display data abstracted from a 24-hour fish recall.)

Table 7-1. Example of 24-hour fish recall data, ounces

	First 24-hour recall				Second 24-hour recall			
	Not locally caught		Locally caught		Not locally caught		Locally caught	
Participant	Tuna	Crab	Bass	Walleye	Tuna	Crab	Bass	Walleye
1	6.5	0	0	0	0	0	4.0	0
2	0	3.0	4.5	0	6.0	0	0	8.5
3	0	8.0	0	3.0	0	0	0	5.5
4	0	0	0	0	0	0	0	0
5	0	0	6.0	0	---	---	---	---

For the desired analysis objective (UFCR of locally caught fish), these data would then be processed to sum up the ounces of reported locally caught fish for each participant-recall day, as shown in Table 7-2.

Table 7-2. Example of summary of 24-hour fish recall data to derive analysis variables, ounces

	First 24-hour recall	Second 24-hour recall
Participant	Locally caught	Locally caught
1	0	4.0
2	4.5	8.5
3	3.0	5.5
4	0	0
5	6.0	---

The general format of the analytic data set is one record per person-recall. Thus, for our example, each participant will have two records in the data set, as shown in Table 7-3. Note that if some participants did not provide a second recall (lost to followup, refused, etc.), as is the case with

participant 5 in our example, the data from their first contact can still be used in the analysis. Ancillary data will be appended to the data set for each record.

Table 7-3. Example format of analytic data set

Participant	Recall Day	Grams Fish	Weekend	Season	Gender	Age	Body weight (lbs.)	Fish Freq	Seq
1	1	0	0	Summer	M	35	197	12	0
1	2	113.4	1	Summer	M	35	197	12	1
2	1	127.6	0	Fall	F	22	150	6	0
2	2	241.0	1	Fall	F	22	150	6	1
3	1	85.0	0	Spring	F	51	168	9	0
3	2	155.9	1	Spring	F	51	168	9	1
4	1	0	1	Winter	M	72	164	3	0
4	2	0	0	Winter	M	72	164	3	1
5	1	170.1	0	Summer	F	46	133	8	0

The column headers represent the variables to be used in the analysis.

- "Recall Day" is the first or second 24-hour fish recall
- "Grams fish" is the number of ounces of fish reported consumed on each recall day converted to grams
- "Weekend" is an indicator variable that denotes whether the recall was for a weekend day (Friday, Saturday, or Sunday) or a weekday (Monday, Tuesday, Wednesday, or Thursday), with a "1" indicating weekend and a "0" indicating weekday
- "Season" is the season during which the recalls occurred
- "Gender" is the gender of the participant
- "Age" is the age of the participant
- "Body weight" is the reported weight of the participant (in this example it is reported in pounds, but the pounds could later be converted to kilograms if necessary.)
- "Fish Freq" is the frequency that fish is consumed, as reported by the participant (This could be the reported frequency for the season during which a participant is interviewed, or it could be the reported frequency for 30 days. It obviously depends on how the questions are asked. Ideally, they will ask about the season during which the interview is being conducted.)
- "Seq" is a variable used in the MIXTRAN macro that accounts for effects due to the sequence number of a subject's records (Since our example data has two records per person, we will need just one seq indicator variable, set to 0 for the person's first record and set to 1 for the second record. If we collected three fish recalls per person, then two indicator variables would be used to identify a subject's second and third records, respectively.)

Note that other ancillary data could also be used (e.g., race and ethnicity).

7.3.2 Application of the NCI Method

The following is a brief discussion of the MIXTRAN and DISTRIB macros. Note that the use of these macros requires knowledge of SAS software and a familiarity with nonlinear mixed modeling for implementation. The macros and more detailed information, a user's guide, and examples can be found on NCI's website at: http://epi.grants.cancer.gov/diet/usualintakes/macros.html .

> "The MIXTRAN macro is used for the analysis of episodically consumed foods, foods consumed every day, and nutrients, and output from the MIXTRAN macro is used by the DISTRIB macro for estimation of the distribution of usual intake. For episodically consumed foods, the MIXTRAN macro fits a two-part nonlinear mixed model where the first part considers the probability of consumption and the second part considers the consumption-day amount. The model allows for covariates and includes a random effect in both parts and allows for correlation between the random effects (Tooze et al., 2006). To fit this nonlinear mixed model with correlated random effects (i.e., the correlated model), starting values for the two parts of the model are obtained by first using the GENMOD procedure to fit a probability model and an amount model. Then a nonlinear mixed model with uncorrelated random effects (i.e., the uncorrelated model) is fit using two calls to the NLMIXED procedure, and the parameter estimates from this model are used as starting values for the correlated model." (NCI Method, MIXTRAN macro, version 2.1)

> "The DISTRIB macro uses results from the MIXTRAN macro and estimates the distribution of usual intake for episodically consumed foods, foods consumed every day, and nutrients (Tooze et al., 2006). The data can then be used to calculate percentiles, and optionally, the percent meeting or failing to meet the recommended daily intake for a population. The DISTRIB macro contains two main functions. First, the DISTRIB macro reads data sets of parameter estimates predicted values output by the MIXTRAN macro. Monte Carlo simulation of the random effect(s) is used to estimate the distribution of usual intake. This data set can be saved. Second, once the data containing the estimated usual intake are available, percentiles and cutpoints can be calculated. The addition of a subgroup variable is accommodated, so that statistics can be calculated by subgroup and for the overall data set. Optionally the percent who meet recommended daily intake values can be calculated. To accomplish this and allow flexibility, the DISTRIB macro contains two sub-macros and some general code to set up and call the macros as requested. The macro MC uses Monte Carlo simulation of the random effect(s) to estimate the distribution of usual intake. The output data set can be saved for future use. The macro PC reads in the usual intake values calculated in the macro MC, normalizes the weights, calculates the percentiles of usual intake, cutpoints if requested, and optionally the percent meeting the recommended intake. A single subgroup variable can be accommodated by the macro PC. The

> resulting data set can be saved for future use.." (NCI Method, DISTRIB macro, version 2.1)

In some rare situations in which a study with weighted data and a large number of primary sampling units is interested in estimating FCR for a number of different fish types (e.g., freshwater fish, estuarine fish, marine, fish for specific trophic levels, and any combination of these) and a number of different populations (e.g., women of childbearing age, children of various age groups, different racial and ethnic groups, varying income levels, etc.), the NCI Method may be too time-consuming (in terms of computer time) to be of practical use. Additionally, the NCI Method has a maximum number of predictors that can be included in the model. Thus if a study finds that there are many statistically significant interactions of important predictors, this may also preclude the use of the NCI Method. Please see U.S. EPA, 2014, for further information and an alternative approach which approximates the NCI Method. Additionally, the MSM, the Iowa State University Method, and SPADE may also be considered.

8 SUMMARY AND APPENDIX OVERVIEW

The main body of this document presents a variety of survey approaches and other considerations for conducting a fish consumption survey. Tables 3-5 through 3-7 present a graphical overview of the various types of sampling frames and modes that can be used, including the strengths and weaknesses of each approach. Choosing an appropriate approach must take into consideration (1) the objectives of the survey, (2) the population being surveyed, and (3) the resources available for the survey. In addition to selecting a sampling frame and mode to satisfy the goals of the planned research, a variety of factors such as characteristics of the population (e.g., literacy rates) and desired outcomes of the survey (e.g., response rates) also factor into the decision. During the planning of the survey, the trade-offs between data desires, data needs, data quality, survey length, representativeness, survey cost, and other factors must be considered. Usually, one or more of these factors will limit and/or more clearly define what the survey can expect to accomplish. Understanding these factors, and their impact on the feasibility of an approach can assist in the selection of the best survey approach to meet research needs.

This document provides guidance on dietary collection methodologies using the approaches summarized in Tables 3-5 through 3-7. Specifically, the reader is provided with a description of an existing instrument, EPA's Fish Consumption Survey Tool (FCST), that can be used to collect data on fish consumption. Various factors that are relevant for instrument development, such as respondent recall periods, use of visual aids, portion sizes, etc., are also discussed. The use of qualitative methods and other pilot testing that can be used to finalize a survey instrument are described.

For some populations or data uses, a typical survey approach may not be appropriate due to suppression of fish consumption. For this reason, this document provides guidance on estimating the historic or heritage FCR by recreating the historic, ecological conditions through the use of historical records and documents and applying knowledge of nutrition, natural resource use, and activity levels. In many cases, heritage rates may be the only practical way to estimate unsuppressed rates – that is, free from the biasing influence of suppression effects. While this approach could be particularly useful for tribal fishing populations, it may not be feasible for other populations or situations where there is a lack of necessary historical information or other data. Therefore, alternative methodologies for establishing unsuppressed rates are presented.

To support the analysis of collected survey data, Table 4-1 presents the data needs, advantages, and limitations of various analytical methodologies for estimating a usual fish consumption rate (UFCR).

In the following appendices (A-D), there are useful tools, materials, and information to assist the reader with the application of guidance presented in the main body of this document. Appendix A includes a discussion of additional considerations for instrument development to help the researcher develop questions that are methodologically sound, and provides example questions for use in existing surveys. Appendix B presents background on a variety of generally accepted dietary assessment methods. Appendix C provides details for the application of the NCI Method and Appendix D provides documentation for the publicly available, automated FCST. This includes a hard-copy version of the instrument, guidance for developing a booklet of photographs to assist respondents in accurately identifying fish species and portion sizes, a supervisor's guide to setting up and configuring the automated tool, and an interviewer's guide.

The appendices, in combination with the main body of this guidance document, provide the reader with an overall strong foundation from which to begin the process of designing, executing, and analyzing a fish consumption survey tailored to meet specific research needs.

9 REFERENCES

Ahmed, F.E., Hattis, D., Wolke, R.E., and Steinman, D. (1993). Human health risks due to consumption of chemically contaminated fishery products. *Environmental Health Perspectives,* 101(Suppl. 3): 297-302.

Andrews, W.R., Brick, J.M., Mathiowetz, N., and Stokes, L. (2010). *Pilot test of a dual frame two-phase mail survey of anglers in North Carolina. Final report*. Prepared for the National Marine Fisheries Service, National Oceanic and Atmospheric Administration, November 2010.

Ashford, J.R., Jones, C.M., and Fegley, L. (2009). A license registry improves sampling efficiency for a marine recreational survey. *Transactions of the American Fisheries Society,* 138(5), 984-989.

Beehler, G.P., McGuinness, B.M., and Vena, J.E. (2001). Polluted fish, sources of knowledge, and the perception of risk: Contextualizing African American anglers' sport fishing practices. *Human Organization,* 60(3), 288-297.

Birch, R.J., Bigler, J., Rogers, J., Zhuang, Y., and Clickner, R.B. (2014). Trends in blood mercury concentrations and fish consumption among U.S. women of reproductive age, NHANES, 1999-2010. *Environmental Research.* 131:431-8.

Blumberg, S.J. and Luke J.V. (2014). *Wireless substitution: Early release of estimates from the National Health Interview Survey, January–June 2014.* National Center for Health Statistics. Available from: http://www.cdc.gov/nchs/nhis/index.htm (accessed December 14, 2014).

Bogen, K. (1996). The effect of questionnaire length on response rates: A review of the literature. In *Proceedings of the Section on Survey Research Methods* (pp. 1020-1025).

Brick, J. M., Brick, P. D., Dipko, S., Presser, S., Tucker, C., & Yuan, Y. (2007). Cell phone survey feasibility in the US: Sampling and calling cell numbers versus landline numbers. *Public Opinion Quarterly*, 71(1), 23-39.

Brick, J.M., Andrews, W.R., Brick, P.D., King, H., Mathiowetz, N.A., and Stokes, L. (2012). Methods for improving response rates in two-phase mail surveys. *Survey Practice,* 5(3).

Brick, J.M., and Williams, D. (2013). Explaining rising nonresponse rates in cross-sectional surveys. *The ANNALS of the American Academy of Political and Social Science,* 645(1), 36-59.

Burger, J., Stephens, W.L., Boring, S., Kuklinski, M., Gibbons, J.W., Gochfeld, M. (1999). Factors in exposure assessment: ethnic and socioeconomic difference in fishing and consumption of fish caught along the Savannah River. *Risk Analysis*;19:427–38.

Carriger, J.F. and Barron, M.G. (2011). Minimizing risks from spilled oil to ecosystem services using influence diagrams: the Deepwater Horizon spill response. *Environmental Science & Technology.* 45(18):7631-9.

Carroll, R.J., Midthune, D., Subar, A.F., Shumakovich, M., Freedman, L.S., Thompson, F.E., Kipnis, V. (2012). Taking advantage of the strengths of 2 different dietary assessment instruments to improve intake estimates for nutritional epidemiology. *American Journal of Epidemiology*. 2012 Feb 15;175(4):340-7.

Connelly, N.A., Brown, T.L., and Decker, D.J. (2003). Factors affecting response rates to natural resource-focused mail surveys: empirical evidence of declining rates over time. *Society and Natural Resources,* 16(6), 541-549.

Connelly, N.A., Lauber, T.B., Niederdeppe, J., and Knuth, B.A. (2012). *Factors affecting fish consumption among licensed anglers living in the Great Lakes Region.* HDRU Publ. No. 12-3. Dept. of Nat. Resour., N.Y.S. Coll. Agric. and Life Sci., Cornell Univ., Ithaca, N.Y. 78 pp.

Connelly, N.A., Smith, K.K., Lauber, T.B., Niederdeppe, J., and Knuth, B.A. (2013). *Factors affecting fish consumption among new mothers living in Minnesota, Pennsylvania, and Wisconsin.* Human Dimensions Research Unit Series Publication No. 13-1. Dept. of Natural Resources, Cornell University, Ithaca, NY.

Connor, G.S., Tremblay, M., Moher, D., and Gorber, B. (2007). A comparison of direct vs. self-report measures for assessing height, weight, and body mass index: A systematic review. *Obesity Reviews.* 8(4): 307-36.

CRITFC (Columbia River Inter-Tribal Fish Commission) (1994). A fish consumption survey of the Umatilla, Nez Perce, Yakama, and Warm Springs Tribes of the Columbia River Basin. *CRITFC Technical Report 94-3*. Portland, Oregon.

Dekkers, A.L., Verkaik-Kloosterman, J., van Rossum, C.T., and Ocké, M.C. (2014). SPADE, a new statistical program to estimate habitual dietary intake from multiple food sources and dietary supplements. *The Journal of Nutrition,* Dec;144(12):2083-91.

de Leeuw, E.D. (2005). To mix or not to mix data collection modes in surveys. *Journal of Official Statistics,* 21(2), 233.

de Leeuw, E.D., Callegaro, M., Hox, J., Korendijk, E., and Lensvelt-Mulders, G. (2007). The Influence of Advance Letters on Response in Telephone Surveys: A Meta-Analysis. *Public Opinion Quarterly,* 71(3), 413-443.

Dietary Assessment Primer. National Institutes of Health, National Cancer Institute (2014). (https://dietassessmentprimer.cancer.gov/) [Accessed on: November 12, 2014].

Dillman, D.A., Smyth, J.D., and Christian, L.M. (2009). *Internet, mail, and mixed-mode surveys: The tailored design method.* Hoboken, NJ: Wiley.

Donatuto, J., and Harper, B.L. (2008). Issues in evaluating fish consumption rates for Native American Tribes. *Risk Analysis,* 28(6), 1497-1506.

Ellis, J.H., Friedman, D.B., Puett, R., Scott, G.I., and Porter, D.E. (2014). A qualitative exploration of fishing and fish consumption in the Gullah/Geechee culture. *J Community Health.* Dec; 39(6):1161-70.

Florida Fish and Wildlife Conservation Commission. (2014). http://myfwc.com/research/freshwater/sport-fishes/management-resources/freshwater-angler-surveys/ Accessed on December 18, 2014

Forsberg, N.D., Stone, D., Harding, A.K., Harper, B., Harris, S., Matzke, M.M., Cardenas, A., Waters, K.M., Anderson, K.A. (2012). Effect of Native American fish smoking methods on dietary exposure to polycyclic aromatic hydrocarbons and possible risks to human health. *Journal of Agricultural and Food Chemistry,* 60(27), 6899-6906.

Fowler, F.J. (1995). *Improving survey questions: Design and evaluation (Applied Social Research Methods Series, Volume 38)*. Thousand Oaks, CA: Sage Publications.

Freedman, L.S., Midthune, D., Carroll, R.J., Krebs-Smith, S., Subar, A.F., Troiano, R.P., Dodd, K., Schatzkin, A., Ferrari, P., and Kipnis, V. (2004). Adjustments to improve the estimation of usual dietary intake distributions in the population. *Journal of Nutrition*.134:1836-1843.

Freedman, L.S., Guenther, P.M., Dodd, K.W., Krebs-Smith, S.M., Midthune, D. (2010). The population distribution of ratios of usual intakes of dietary components that are consumed every day can be estimated from repeated 24-hour recalls. *The Journal of Nutrition,* Jan; 140(1):111-6.

Gaziano, C. (2005). Comparative analysis of within-household respondent selection techniques. *Public Opinion Quarterly,* 69(1), 124-157.

Gersovitz, M., Madden, J.P., Smiciklas-Wright, H. (1978). Validity of the 24-hour dietary recall and seven-day record for group comparisons. *Journal of the American Dietetic Association,* 73, 48-55.

Goldstein, K.M., and Jennings, M.K. (2002). The effect of advance letters on cooperation in a list sample telephone survey. *Public Opinion Quarterly,* 66(4), 608-617.

Groves, R.M. (2006). Nonresponse rates and nonresponse bias in household surveys. *Public Opinion Quarterly,* 70(5), 646-675.

Groves, R.M., Cialdini, R.B., and Couper, M.P. (1992). Understanding the decision to participate in a survey. *Public Opinion Quarterly,* 56(4), 475-495.

Haines, P.S., Hama, M.Y., Guilkey, D.K., and Popkin, B.M. (2003). Weekend eating in the United States is linked with greater energy, fat, and alcohol intake. *Obesity Research,* 11: 945–949.

Hansen, K.M. (2007). The effects of incentives, interview length, and interviewer characteristics on response rates in a CATI-study. *International Journal of Public Opinion Research,* 19(1), 112-121.

Harding, A., Harper, B., Stone, D., O'Neill, C., Berger, P., Harris, S., and Donatuto, J. (2012). Conducting research with tribal communities: sovereignty, ethics, and data-sharing issues. *Environmental Health Perspectives,* 120(1), 6.

Harkness, J.A., Villar, A., and Edwards, B. (2010). Translation, Adaptation, and Design. In *Survey methods in multinational, multiregional, and multicultural contexts*. Harkness, J.A., Braun, M.,

Edwards, B., Johnson, T.P., Lyber, L., Mohler, P., Pennell, B. E., and Smith, T.W. (Eds.). Hoboken, NJ: John Wiley and Sons.

Harper, B., Flett, B., Harris, S., Abeyta, C., Kirschner, F. (2002). The Spokane Tribe's multipathway subsistence exposure scenario and screening level RME. *Risk Analysis,* 22(3):513-26.

Harper, B. (2005). *Washoe Tribe Human Health Risk Assessment Exposure Scenario for the Leviathan Mine Superfund Site.* Available at: http://health.oregonstate.edu/sites/default/files/research/pdf/tribal-grant/Washoe-HHRA-Exposure-Scenario-FINAL-3-17-05.pdf [accessed 16 April 2015].

Harper, B.L., Harding, A.K., Waterhous, T., Harris, S.G. (2007). Traditional Tribal Subsistence Exposure Scenario and Risk Assessment Guidance Manual. Corvallis: Oregon State University. http://superfund.oregonstate.edu/conducting-research-tribal-communities

Harper, B. (2008). *Quapaw Traditional Lifeways Exposure.* AESE, Inc; Richland, WA, USA: 2008. Available at: http://superfund.oregonstate.edu/sites/default/files/harper_2008_quapaw_scenario_final.pdf [accessed 16 April 2015].

Harper, B., and Ranco, D. (2009). *Wabanaki traditional cultural lifeways exposure scenario.* Available at: http://superfund.oregonstate.edu/sites/superfund.oregonstate.edu/files/harper_ranco_2009_wabanaki_scenario_final.pdf[accessed 3 December 2014].

Harper, B., Harding, A.K., Harris, S., Berger, P. (2012). Subsistence exposure scenarios for tribal applications. *Human and Ecological Risk Assessment: An International Journal,* 18(4):810-831.

Harper, B., Walker, D.E. (2015a). Comparison of Contemporary and Heritage Fish Consumption Rates in the Columbia River Basin. Human Ecology: Volume 43, Issue 2. Page 225-236. DOI: 10.1007/s10745-015-9734-4

Harper, B. and Walker, D.E. (2015b). Columbia Basin Heritage Fish Consumption Rates. Human Ecology: Volume 43, Issue 2. Page 237-245. DOI: 10.1007/s10745-015-9735-3

Harris, S.G. and Harper, B.L. (2004) *Exposure Scenario for CTUIR Traditional Subsistence Lifeways.* Department of Science & Engineering, Confederated Tribes of the Umatilla Indian Reservation, P.O. Box 638, Pendleton, OR 97801. Available at: http://health.oregonstate.edu/sites/default/files/research/pdf/tribal-grant/CTUIR-SCENARIO.pdf [accessed 16 April 2015].

Harttig, U., Haubrock, J., Knüppel, S., Boeing, H.; EFCOVAL Consortium (2011). The MSM program: Web-based statistics package for estimating usual dietary intake using the Multiple Source Method. *European Journal of Clinical Nutrition,* 2011 Jul; 65 Suppl 1:S87-91. doi: 10.1038/ejcn.2011.92.

Hill, D.H., and Willis, R.J. (2001). Reducing panel attrition: A search for effective policy instruments. *Journal of Human Resources,* 36(3), 416-438.

Holbrook, A.L., Green, M.C., and Krosnick, J.A. (2003). Telephone versus face-to-face interviewing of national probability samples with long questionnaires: Comparisons of respondent satisficing and social desirability response bias. *Public Opinion Quarterly,* 67(1), 79-125.

Holt, D. and Smith, T.M.F. (1979). Post stratification. *Journal of the Royal Statistical Society*, Series A, 142, Part 1, 33-46.

Idler, D.R., and W.A. Clemens. (1959). *The Energy Expenditures of Fraser River Sockeye Salmon During the Spawning Migration to Chilko and Stuart Lakes.* Intern. Pacific Salmon Fish. Comm., Prog. Rept. No. 6, New Westminster, B.C.).

Imm P., Knobeloch, L., Anderson, H.A., and the Great Lakes Sport Fish Consortium. (2005). Fish consumption and advisory awareness in the Great Lakes Basin. *Environmental Health Perspectives*, 113(10): 1325-1329.

Israel, B.A., Schulz, A.J., Parker, E.A., Becker, A.B. (1998). Review of community-based research: assessing partnership approaches to improve public health. *Annual Review of Public Health*. 19:173–202.

Israel, B.A., Eng, E, Schulz, A.J., Parker, E.A., editors. (2012). *Methods in Community-Based Participatory Research for Health*, (2nd Edition). San Francisco, CA: Jossey-Bass.

Judd, N.L., Griffith, W.C., Faustman, E.M. (2004). Consideration of cultural and lifestyle factors in defining susceptible populations for environmental disease. *Toxicology,* 198(1-3):121-33.

Katner, A., Ogunyinka, E., Sun, M.H., Soileau, S., Lavergne, D., Dugas, D., and Suffet, M. (2011). Fishing, fish consumption and advisory awareness among Louisiana's recreational fishers. *Environmental Research,* 111(8), 1037-1045.

Kinnell, J.C., Bingham, M.F., Hastings, E.A., Ray, R., Craven, V., and Freeman, M. (2007). A survey methodology for collecting fish consumption data in urban and industrial water bodies (part 1). *Journal of Toxicology and Environmental Health, Part A,* 70(6), 477-495.

Kipnis, V., Subar, A.F., Midthune, D., Freedman, L.S., Ballard-Barbash, R., Troiano, R.P., Bingham, S., Schoeller, D.A., Schatzkin, A., and Carroll, R.J. (2003). Structure of dietary measurement error: Results of the OPEN biomarker study. *American Journal of Epidemiology*. 158:14-21.

Kipnis, V., Midthune, D., Buckman, D.W., Dodd, K.W., Guenther, P.M., Krebs-Smith, S.M., Subar, A.F., Tooze, J.A., Carroll, R.J., Freedman, L.S. (2009). Modeling data with excess zeros and measurement error: application to evaluating relationships between episodically consumed foods and health outcomes. *Biometrics,* Dec; 65(4):1003-10.

Kissinger, L., Lorenzana, R., Mittl, B., Lasrado, M., Iwenofu, S., Olivo, V., Helba, C., Capoeman, P., Williams, A.H. (2010). Development of a computer-assisted personal interview software system for collection of tribal fish consumption data. *Risk Analysis.* Dec; 30(12):1833-41. doi: 10.1111/j.1539-6924.2010.01461.x.

Kroke, A., Klipstein-Grobusch, K., Voss, S., Möseneder, J., Thielecke, F., Noack, R., and Boeing, H. (1999). Validation of a self-administered food-frequency questionnaire administered in the

European Prospective Investigation into Cancer and Nutrition (EPIC) Study: comparison of energy, protein, and macronutrient intakes estimated with the doubly labeled water, urinary nitrogen, and repeated 24-h dietary recall methods. *American Journal of Clinical Nutrition,* Oct;70(4):439-47.

Krosnick, J.A. (1991). Response strategies for coping with the cognitive demands of attitude measures in surveys. *Applied Cognitive Psychology,* 5, 213-236.

Link, M.W., and Mokdad, A. (2005). Advance letters as a means of improving respondent cooperation in random digit dialing studies: A multistate experiment. *Public Opinion Quarterly,* 69(4), 572-587.

Link, M. W., Battaglia, M. P., Frankel, M. R., Osborn, L., & Mokdad, A. H. (2007). Reaching the US cell phone generation comparison of cell phone survey results with an ongoing landline telephone survey. *Public Opinion Quarterly*, 71(5), 814-839.

Lohr, S. (2010). *Sampling: Design and analysis,* (2nd edition). Boston, MA: Brooks/Cole, Cengage Learning.

Mahaffey, K.R., Clickner, R.P., Jeffries, R.A. (2009). Adult women's blood mercury concentrations vary regionally in the United States: Association with patterns of fish consumption (NHANES 1999-2004). *Environmental Health Perspectives,* 117(1), 47-53.

Mercer, A., Caporaso, A., Cantor, D., Townsend, R. (2015). How Much Gets You How Much? Monetary Incentives and Response Rates in Household Surveys. *Public Opinion Quarterly*, 79(1), Spring 2015, 105-129.

Merrill, R.M. and Richardson, J.S. (2009). Validity of self-reported height, weight, and body mass index: Findings from the National Health and Nutrition Examination Survey, 2001-2006. *Preventing Chronic Disease,* 6(4). http://www.cdc.gov/pcd/issues/2009/oct/08_0229.htm Accessed December 18, 2014.

Michigan Department of Community Health. (2007). California Recreational Fisheries Survey. Available at: http://www.michigan.gov/documents/mdch/FCS_Final_rpt_061407_199288_7.pdf. Last accessed January 15, 2015.

Michigan Department of Natural Resources. (2014).http://www.michigan.gov/dnr/0,4570,7-153-10364_52259_10951_11301-97829--,00.html.

Millar, M.M., and Dillman, D.A. (2011). Improving response to web and mixed-mode surveys. *Public Opinion Quarterly,* 75(2), 249-269.

Minelli, C. and Baio, G. (2015). Value of Information: A Tool to Improve Research Prioritization and Reduce Waste. PLoS Medicine, 12(9).

Minkler. M, and Wallerstein, N., editors. (2003). *Community-Based Participatory Research for Health*. San Francisco: Jossey-Bass Publishers.

Minnesota Department of Health. (2012). *Fish consumption and fish advisory awareness among Minnesota women who recently gave birth*. Minnesota Department of Health Environmental Division, Environmental Surveillance and Assessment Section, Site Assessment and Consultation Unit. Available at: http://www.health.state.mn.us/divs/eh/fish/techinfo/momsurvey1.pdf

Minnesota Department of Natural Resources. (2011). http://www.dnr.state.mn.us/areas/fisheries/montrose/creelsurvey.html

National Cancer Institute, Register of Validated Short Dietary Assessment Instruments. (2013). Applied Research Program, National Cancer Institute, http://appliedresearch.cancer.gov/diet/shortreg/.

National Center for Education Statistics. (2003). U.S. Department of Education. *National assessment of adult literacy*. Available at: http://nces.ed.gov/naal/index.asp. Last accessed December 3, 2014.

National Environmental Justice Advisory Council (NEJAC). Fish Consumption and Environmental Justice; 2002. Available at: https://www.epa.gov/environmentaljustice/epa-fish-consumption-and-environmental-justice

National Research Council. (2000). *Committee on the toxicological effects of methylmercury*. Washington, DC: National Academies Press.

New Jersey Department of Environmental Protection. (2014). Division of Fish and Wildlife. http://www.njfishandwildlife.com/marinesurvey.htm Accessed December 18, 2014.

Nusser, S.M., Carriquiry, A.L., Dodd, K.W., and Fuller, W.A. (1996). A semi-parametric transformation approach to estimating usual daily intake distributions. *Journal of the American Statistical Association,* 91, 1440-1449.

Office of Management and Budget. 2006. "Standards and Guidelines for Statistical Surveys." Available at http://www.whitehouse.gov/sites/default/files/omb/inforeg/statpolicy/standards_stat_surveys.pdf

Oh, H. and Scheuren F. (1983). Weighting adjustments for unit nonresponse. *Incomplete Data in Sample Surveys, Volume 2: Theory and Bibliographies* (W. Madow, I. Olkin and D. Rubin, eds), 143-184. New York: Academic Press.

O'Neill, C. Fishable waters, *1. American Indian Law Journal,* 181 (2013). Available at: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2263490.

Oregon Department of Fish and Wildlife. (2014). https://nrimp.dfw.state.or.us/crl/default.aspx?pn=BACS Accessed December 18, 2014.

Parsons, G.R., Morgan, A., Whitehead, J.C., and Haab, T.C. (2006). The welfare effects of pfiesteria-related fish kills: A contingent behavior analysis of seafood consumers. *Agricultural and Resource Economics Review,* 35.2: 349-356.

Pew Internet and American Life Project. (2013).

Ray, R., Craven, V., Kinnell, J., Bingham, M., Freeman, M., and Finley, B. (2007). A statistical method for analyzing data collected by a creel/angler survey (part 2). *Journal of Toxicology and Environmental Health. Part A.* Mar 15; 70(6):496-511.

Ray, R., Craven, V., Bingham, M., Kinnell, J., Hastings, E., and Finley, B. (2007). Human health exposure factor estimates based upon a creel/angler survey of the lower Passaic River (part 3). *Journal of Toxicology and Environmental Health. Part A.* Mar 15; 70(6):512-28.

Razzaghi, H., and Tinker, S.C. (2014). Seafood consumption among pregnant and non-pregnant women of childbearing age in the United States, NHANES 1999-2006. *Food & Nutrition Research.* Jun 11;58.

Salama, A.A., Mohamed, M.A.M., Duval, B., Potter, T.L., and Levin, R.E. (1998). Polychlorinated biphenyl concentration in raw and cooked north Atlantic bluefish fillets. *Journal of Agricultural and Food Chemistry,* 46:1359–1362.

Skea, J.C., Simonin, H.A., Harris, E.J., Jackling, S., Spagnoli, J.J., Symula, J., et al. (1979). Reducing levels of mirex, Aroclor 1254, and DDE by trimming and cooking Lake Ontario brown trout and smallmouth bass. *Journal of Great Lakes Research,* 5(2):153–159.

Soon, R., Dye, T.D., Ralston, N.V., Berry, M.J, and Sauvage, L.M. (2014). Seafood consumption and umbilical cord blood mercury concentrations in a multiethnic maternal and child health cohort. *BMC Pregnancy & Childbirth.* Jun 18;14:209.

Souverein, O.W., Dekkers, A.L., Geelen, A., Haubrock, J., de Vries, J.H., Ocké, M.C., Harttig, U., Boeing, H., van 't Veer, P. (2011). Comparing four methods to estimate usual intake distributions. *European Journal of Clinical Nutrition,* Jul;65 Suppl 1:S92-101.

Stahl, L.L., Snyder, B.D., Olsen, A.R., and Walters, L.S. (2013). A national probabilistic study of polybrominated diphenyl ethers in fish from US lakes and reservoirs. *Environmental Monitoring and Assessment,* 185:10351-10364.

Stahl, L.L., Snyder, B.D., Olsen, A.R., Kincaid, T.M., Wathen, J.B., and McCarty, H.B. (2014). Perfluorinated compounds in fish from U.S. urban rivers and the Great Lakes. *Science of the Total Environment,* 499:185-195.

Subar, A.F., Thompson, F.E., Kipnis, V., Midthune, D., Hurwitz, P., McNutt, S., McIntosh, A., and Rosenfeld, S. (2001). Comparative validation of the Block, Willett, and National Cancer Institute food frequency questionnaires: the Eating at America's Table Study. *American Journal of Epidemiology,* 154:1089-99.

Subar, A.F., Kipnis, V., Troiano, R.P., Midthune, D., Schoeller, D.A., Bingham, S., Sharbaugh, C.O., Trabulsi, J., Runswick, S., Ballard-Barbash, R., Sunshine, J., and Schatzkin, A. (2003). Using intake biomarkers to evaluate the extent of dietary misreporting in a large sample of adults: The OPEN study. *American Journal of Epidemiology*.158:1-13.

The Suquamish Tribe. (2000). Fish Consumption Survey of the Suquamish Indian Tribe of the Port Madison Indian Reservation. Puget Sound Region. August 2000. Report writing group: M. Duncan, N.L. Polissar, S. Liao, D. LaFlamme.

Tarrant, M.A., and Manfredo, M.J. (1993). Digit preference, recall bias, and nonresponse bias in self reports of angling participation. *Leisure Sciences,* 15, 231–238.

Tennessee Wildlife Resources Agency. (2014). http://www.tnfish.org/TournamentFishingTennessee_TWRA/TWRA_BiteProgram.htm. Accessed on December 18, 2014.

Texas Parks and Wildlife Department. (2014). https://www.tpwd.state.tx.us/fishboat/fish/didyouknow/creel.phtml. Accessed on December 18, 2014.

Thompson, F.E., Subar, A.F. (2008). Dietary assessment methodology. In: Coulston AM, Rock CL, Monsen ER, eds. Nutrition in the prevention and treatment of disease. San Diego (CA): Academic Press; 2008.

Thompson, M.R., and Boekelheide, K. (2013). Multiple environmental chemical exposures to lead, mercury and polychlorinated biphenyls among childbearing-aged women (NHANES 1999-2004): Body burden and risk factors. *Environmental Research,* 2013 Feb;121:23-30. doi: 10.1016/j.envres.2012.10.005. Epub 2012 Nov 16.

Tooze, J.A., Kipnis, V., Buckman, D.W., Carroll, R.J., Freedman, L.S., Guenther, P.M., Krebs-Smith, S.M., Subar, A.F., Dodd, K.W. (2010). A mixed-effects model approach for estimating the distribution of usual intake of nutrients: The NCI method. *Statistics in Medicine,* 2010 Nov 30;29(27):2857-68.

Tooze, J.A., Midthune, D., Dodd, K.W., Freedman, L.S., Krebs-Smith, S.M., Subar, A.F., Guenther, P.M., Carroll, R.J., Kipnis, V. (2006). A new statistical method for estimating the usual intake of episodically consumed foods with application to their distribution. *Journal of the American Dietetic Association,* Oct;106(10):1575-87.

Tourangeau, R., Edwards, B., Johnson, T.P., Wolter, K.M., and Bates, N. (Eds.). (2014). *Hard-to-survey populations.* Cambridge University Press, United Kingdom.

Trondsen, T., Braaten, T., Lund, E., and Eggen, A. E. (2004). Consumption of seafood—the influence of overweight and health beliefs. *Food Quality and Preference,* 15(4), 361-374.

USDA. (2010). Nutrient Data Laboratory Home Page. Agricultural Research Service, U.S. Department of Agriculture. Available at http://www.ars.usda.gov/main/site_main.htm?modecode=12-35-45-00.

USDA. (2014). Food and Nutrient Database for Dietary Studies. Agricultural Research Service, U.S. Department of Agriculture. Available at http://www.ars.usda.gov/services/docs.htm?docID=12089.

U.S. EPA. (1994). *Guidance for the Data Quality Objectives Process.* EPA-600-R-96-055. U.S. Environmental Protection Agency, Office of Research and Development, Washington, DC. Available at https://archive.epa.gov/epawaste/hazard/web/pdf/epaqag4.pdf.

U.S. EPA. (1995). *Guidance for Assessing Chemical Contaminant Data for Use in Fish Advisories. Volume IV. Risk Communication. 3rd edition.* EPA-823-R-95-001. U.S. Environmental Protection Agency, Office of Water, Washington, DC.

U.S. EPA. (1998). *Guidance for Conducting Fish and Wildlife Consumption Surveys.* EPA-823-B-98-007. U.S. Environmental Protection Agency, Office of Water, Washington, DC.

U.S. EPA. (1999). *Asian & Pacific Islander Seafood Consumption Study in King County, WA.* EPA-910/R99-003. U.S. Environmental Protection Agency, Office of Water, Washington, DC.

U.S. EPA. (2000a). *Guidance for Assessing Chemical Contaminant Data for Use in Fish Advisories. Volume I. Fish Sampling and Analysis. 3rd edition.* EPA-823-B-00-007. U.S. Environmental Protection Agency, Office of Water, Washington, DC.

U.S. EPA. (2000b). *Guidance for Assessing Chemical Contaminant Data for Use in Fish Advisories. Volume II. Risk Assessment and Fish Consumption Limits. 3rd edition.* EPA-823-B-00-008. U.S. Environmental Protection Agency, Office of Water, Washington, DC.

U.S. EPA. (2000c). *Guidance for Assessing Chemical Contaminant Data for Use in Fish Advisories. Volume III. Overview of Risk Management. 2nd edition.* EPA-823-B-00-009. U.S. Environmental Protection Agency, Office of Water, Washington, DC.

U.S. EPA. (2000d). *Methodology for Deriving Ambient Water Quality Criteria for the Protection of Human Health (2000).* EPA-822-B-00-004. Washington, DC: U.S. Environmental Protection Agency, Office of Science and Technology, Office of Water.

U.S. EPA. (2002). *Estimated per capita fish consumption in the United States.* EPA-821- C- 02-003. Washington, DC: U.S. Environmental Protection Agency, Office of Science and Technology, Engineering and Analysis Division, Office of Water.

U.S. EPA. (2007). *Framework for Selecting and Using Tribal Fish and Shellfish Consumption Rates for Risk-Based Decision Making at CERCLA and RCRA Cleanup Sites in Puget Sound and the Strait of Georgia.* U.S. Environmental Protection Agency, Region 10, Office of Environmental Cleanup, Office of Air, Waste and Toxics, Office of Environmental Assessment. Available at: http://yosemite.epa.gov/r10/CLEANUP.NSF/7780249be8f251538825650f0070bd8b/e12918970debc8e488256da6005c428e/$FILE/Tribal%20Shellfish%20Framework.pdf

U.S. EPA. (2009). *The National Study of Chemical Residues in Lake Fish Tissue.* EPA-823-R-09-006. Washington, DC: U.S. Environmental Protection Agency, Office of Water.

U.S. EPA. (2011). *Exposure Factors Handbook: 2011 Edition. National Center for Environmental Assessment*, Washington, DC: U.S. Environmental Protection Agency, Office of Research and Development, EPA/600/R-09/052F. Available from the National Technical Information Service, Springfield, VA, and online at https://cfpub.epa.gov/ncea/risk/recordisplay.cfm?deid=236252 .

U.S. EPA. (2013). *Human Health Ambient Water Quality Criteria and Fish Consumption Rates: Frequently Asked Questions.* Washington, DC: U.S. Environmental Protection Agency, Office of Science

and Technology, Office of Water. Available at: https://www.epa.gov/wqc/human-health-water-quality-criteria

U.S. EPA. (2014). *Estimated Fish Consumption Rates for the U.S. Population and Selected Subpopulations (NHANES 2003-2010).* EPA-820-R-14-002. Washington, DC: U.S. Environmental Protection Agency, Office of Science and Technology, Office of Water.

Vaske, J.J., Beaman, J., Manfredo, M.J., Covey, D., and Knox, R. (1996). Response strategy, recall frame and digit preference in self-reports of angling participation. *Human Dimensions of Wildlife,* 14, 54–68.

Voiland, M.P., Gall, K.L., Lisk, D.J., MacNeill, D.B. (1991). Effectiveness of recommended fat-trimming procedures on the reduction of PCB and mirex levels in brown trout from Lake Ontario. *Journal of Great Lakes Research,* 17(4):454–460.

Washington Department of Fish and Wildlife. (2014). http://wdfw.wa.gov/fishing/reports_plants.html.

Washington State Department of Ecology. (2012). *Supplemental Information to Support the Fish Consumption Rates Technical Support Document.* Olympia, WA: Author. Retrieved from http://www.ecy.wa.gov/programs/tcp/regs/fish/2012/FCR-doc.html.

Wathen, J.B., Lazorchak, J.M. , Olsen, A.R., and Batt, A. (2015). A national statistical survey assessment of mercury concentrations in fillets of fish collected in the U.S. EPA national rivers and streams assessment of the continental USA. *Chemosphere,* 122:52-61.

Wentland, E. J., and Smith, K.W. (1993). *Survey responses: An evaluation of their validity.* San Diego, CA: Academic Press.

Willett, W. (2012), Nutritional Epidemiology. Oxford University Press.

Willis, G.B. (2005). *Cognitive interviewing: A tool for improving questionnaire design.* Thousand Oaks, CA: Sage.

Zabik, M.E., Booren, A., Zabik, M.J., Welch, R., and Humphrey, H. (1996). Pesticide residues, PCBs and PAHs in baked, charbroiled, salt boiled and smoked Great Lakes lake trout. *Food Chemistry,* 55(3):231–239.

Zabik, M.E., Zabik, M.J. (1999). Polychlorinated biphenyls, polybrominated biphenyls, and dioxin reduction during processing/cooking food. *Advances in Experimental Medicine and Biology,* 459:213–231.

Zabik, M.E., Zabik, M.J., Booren, A.M., Nettles, M., Song, J.H., Welch, R., et al. (1995). Pesticides and total polychlorinated biphenyls in Chinook salmon and carp harvested from the Great Lakes: effects of skin-on and skin-off processing and selected cooking methods. *Journal of Agricultural and Food Chemistry,* 43:993–1001.

Zale, A.V., Parrish, D.L., and Sutton, T. M. (2013). *Fisheries Techniques, Third Edition*. Bethesda (Maryland): American Fisheries Society.

Zhang, L.C. (2000). Post-stratification and calibration. *American Statistician*, 54, 178-184.

Appendices

APPENDIX A. ADDITIONAL INSTRUMENT DEVELOPMENT CONSIDERATIONS

For a survey to be readily understood by respondents, the questions must be simple and straightforward while accomplishing study objectives. Additionally, questions should consider cultural aspects, such as possible unique common names or preparation techniques for fish species that are specific to the population of interest. The design of each specific question must consider both sentence structure and wording as well as the mode of administration (see, for example, Dillman & Christian, 2005). The position of questions within the survey is also important. These topics are discussed below. The reader is referred to Bradburn et al., 2004, Groves et al., 2009, Saris and Gallhofer, 2007, Tourangeau et al., 2000, and the references cited therein for more information. Pollock et al., 1994, is a good resource for integrating the varied survey disciplines as they apply to fisher surveys. More recent books on fishery survey methods have not focused as well on the issues of question design as it applies to fisher surveys.

If possible the development of an instrument (or questionnaire) should rely heavily on the use of previously validated questions. If local data are to be compared against other data sets (e.g., state or national data), the questions should be asked in the same way so as to ensure consistency of the data collected. Appendix C provides some examples of questions that can be adapted to obtain data on fish consumption and portion sizes.

Question structure: Four general types of question structure are available (Pollack et al., 1994): (1) open-end questions, (2) closed-end questions with ordered response choices; (3) closed-end questions with unordered response choices, and (4) partially closed-end questions. Open-end questions have no categories from which the respondent can choose; however, interpretation of all but the simplest open-end questions can be quite difficult. It should also be noted that standardized coding of open-end questions for analysis can be quite costly and time consuming. Closed-end questions provide several answer categories, which can be ordered sequentially (e.g., numerically) or unordered. The answers to closed-end questions are easy to summarize quantitatively. Response options must be selected carefully so that the choices are mutually exclusive, inclusive of all reasonable choices, and easy to understand. Categories also may provide cues to aid respondents' recall (Bradburn et al., 2004). Partially closed-end questions often allow an open-ended option such as "other." This option represents a good compromise between open-ended and closed-end structures (Pollack et al., 1994). It also should be noted that standardized coding of the "other" category responses may prove difficult and will require careful review to see if any responses should be "up-coded"—that is, recoded into the closed-end categories provided.

For closed-end questions, the specific ranges for each response alternative can affect the way in which the question is answered, resulting in impression-based responses. Values in the middle range of the scale selected are often assumed by respondents to reflect the "average" or "typical" behavior, whereas the extremes of the scale are assumed to represent the extremes of the distribution (Groves et al., 2009; Schwarz & Hippler, 2004). Thus, respondents will assume that the response scale provided indicates knowledge of the behavior and form an impression based on the scale provided. Respondents will then adjust their response based on their impression of how they compare to the scale provided.

Question wording: The specific wording of questionnaires on fish consumption must be developed very carefully to elicit nonbiased responses. Some guidelines for question wording are listed below (Pollack et al., 1994).

- All alternatives of a multiple-choice question should be given
- As few words as necessary should be used
- The units that apply to each response should be given
- The time frame covered by the survey should be clear
- Only one concept or issue should be addressed by each question

Draft questions should be reviewed carefully for any ambiguities and tested through the use of qualitative testing procedures (Presser et al., 2004). For a review of question wording and design guidelines, the reader is directed to Bradburn et al., 2004.

Question order: Topic sections should be arranged for the convenience of the respondent, not that of the researcher. There is likely a logical order to grouping questions that will aid in respondent recall. The questions should build on each other. For example, rather than asking, "Did you wear your seatbelt the last time you rode in a car?," the following series of questions may be more effective: "When was the last time you rode in a car? How long was the trip? Did you wear your seatbelt?" This type of cognitive design can be very effective in minimizing respondent error and should be used for important questions (Groves et al., 2009).

The first few questions of the survey might be the most critical, particularly for self-administered surveys (e.g., mail or web), since these might determine whether the respondent chooses to complete the questionnaire. Sensitive questions or questions that are difficult to answer should be asked near the end of the interview so as not to threaten the respondent and possibly compromise the rapport between interviewer and respondent. Sensitive questions include demographic questions such as age, income, and education (Bradburn et al., 2004), and questions about whether the fisher has an applicable fishing license or is familiar with a particular advisory or regulation (Pollock et al., 1994).

Question design and presentation: Just as question features, such as, context, response scale, and number of response options can influence the answers provided by respondents, the presentation and design of survey questions can also affect how respondents answer. For self-administered paper surveys or mail surveys, the placement and highlighting of skip instructions can affect errors of omission or commission (Dillman et al., 2009). An error of omission occurs when a respondent skips a question that should have been answered, whereas an error of commission occurs when a respondent answers a question that should have been skipped. Another example of design influences happens when, for long response lists, response options are presented in two columns rather than a single column. The use of two columns will result in more selections from the first column, resulting in bias. Errors associated with question branching or skipping can be greatly reduced by use of web or computer assisted interview software approaches that automate skip patterns.

For web surveys, features such as layout, spacing, and proximity can influence how respondents answer questions. Tourangeau et al., 2013, describe five heuristics, or rules, that respondents employ to interpret response scales, as follows.

- Middle means typical – respondents will use the visual midpoint as the mean or typical value
- Left and top mean first – generally that response options will follow a logical order
- Near means related – relative proximity to other options may imply a relationship between those options
- Similarity in appearance means close in meaning – the use of highlighting, or giving similar appearance to another highlighted option, may make the options seem related
- Up means good – vertical position on the screen may imply value about the item

For response fields requiring specific formats, such as dates, the layout of response fields and labels can influence how likely respondents are to follow directions and provide answers appropriately. For example, in a review of a number of studies looking at the proportion of respondents correctly entering a date, the highest proportion of correctly formatted entries occurred when labels mimicking the expected formatting were used in conjunction with close proximity to the desired entry box (Tourangeau et al., 2013). In summary, the survey designer should be aware of the influences that survey design features can have on respondents' answers.

REFERENCES

Bradburn, N.M., Sudman, S., and Wansink, B. (2004). *Asking questions: The definitive guide to questionnaire design—for market research, political polls, and social and health questionnaires.* San Francisco, CA: Jossey-Bass.

Dillman, D.A. and Christian, L.M. (2005). Survey mode as a source of instability in responses across surveys. *Field Methods,* 17.1: 30-52.

Dillman, D.A., Smyth, J.D., & Christian, L.M. (2009). *Internet, mail, and mixed-mode surveys: The tailored design method.* Hoboken, NJ: Wiley.

Groves, R.M., Fowler, F.J., Couper, M.P., Lepkowski, J.M., Singer, E., and Tourangeau, R. (2009). *Survey methodology.* Hoboken, NJ: John Wiley & Sons.

Pollock, K.H., Jones, C.M., and Brown, T.L. (1994). Angler survey methods and their applications in fisheries management. *American Fisheries Society Special Publication 25,* American Fisheries Society, Bethesda, Maryland.

Presser, S., Rothgeb, J.M., Couper, M.P., Lessler, J.T., Martin, E., Martin, J., and Singer, E. (2004). *Methods for testing and evaluating survey questionnaires. Vol. 546.* Hoboken, NJ: John Wiley & Sons.

Saris, W.E. and Gallhofer, I.N. (2007). *Design, evaluation, and analysis of questionnaires for survey research.* Hoboken, NJ: John Wiley & Sons.

Schwarz, N. and Hippler, H.-J. (2004). *Response alternatives: The impact of their choice and presentation order, in measurement errors in surveys* (eds. P. P. Biemer, R. M. Groves, L. E. Lyberg, N. A. Mathiowetz and S. Sudman), Hoboken, NJ: John Wiley & Sons. doi: 10.1002/9781118150382.ch3

Tourangeau, R., Rips, L.J., and Rasinski, K. (2000). *The psychology of survey response.* Cambridge, U.K.: Cambridge University Press.

Tourangeau, R., Conrad, F., and Couper, M.. (2013). *The science of web surveys.* New York, NY: Oxford University Press.

APPENDIX B. OVERVIEW OF DIETARY ASSESSMENT METHODS

Method	Description
24-hour recall (24HR)	A very structured method in which the respondent is asked, using standard probes, to recall and describe all food and beverages consumed, or only a specific food like fish, in the preceding 24 hours. Can be interviewer- or self-administered either in-person, by telephone, or web. Typically, data are captured using a standard software program. Portion size estimating aids assist the respondent to recall amounts collected. Advantages: Provides the highest quality and least biased food intake data for a single day. Can be used to collect total diet or just fish consumption. Analyzes individual data for grams consumed per food and per day. If interviewer-administered, low respondent burden and few literacy issues. Single day used to estimate average dietary intake of a group; multiple days (or when paired with FFQ data) used to estimate population distributions (NCI method described in Chapter 7). Limitations: Multiple days needed or combined with FFQ to model estimates of usual intake (Tooze et al., 2006; Carroll et al., 2012). A measure of frequency of consumption is needed to improve the statistical modeling. Moderate to high cost depending on mode of administration. Some measurement error due to underreporting; may require trained staff (interviewers and dietary coders) depending on mode of administration. Software programs do not include specific species of finfish and shellfish. Requires respondent to estimate portion of fish consumed. Instruments: • U.S. Department of Agriculture (USDA) Automated Multiple Pass Method (AMPM). Interviewer-administered software used in the National Health and Nutrition Examination Survey (NHANES) to collect surveillance data. Over 2,500 food probes (80 probes related to fish) used to collect very detailed descriptive, source, and portion data. Uses the USDA Food and Nutrient Database for Dietary Studies (FNDDS) to apply food code, portion sizes, and nutrients (includes over 230 primary fish codes and 120 fish dishes). Requires post-interview processing system (PIPS) and SurveyNet coding system to generate a data set with total fish consumption by person. These programs can be requested from USDA A Agricultural Research Service and may require setting up a cooperative agreement. • National Cancer Institute (NCI) Automated Self-Administered 24-hour Recall (ASA24). English and Spanish versions of web-based software that can be self- or interviewer-administered. Uses same food database and detailed probes to collect descriptive, source, and portion data. Validated against the AMPM (Kirkpatrick et al., 2014). Includes 95 finfish and shellfish varieties. All reported foods are auto-coded, so no dietary coding is required. This software is freely available at http://appliedresearch.cancer.gov/diet/usualintakes/macros.html. • University of Minnesota Nutrition Coordinating Center (NCC) Nutrition Data System for Research (NDS-R). Windows-based, interviewer-administered, multiple-pass dietary analysis program designed to collect and analyze total diet

Method	Description
	using the NDS database of over 18,000 foods including 8,000 brand names (over 480 primary fish codes and 240 fish dishes) . This software is available at http://www.ncc.umn.edu/products/ndsr.html and requires a licensing fee.
Fish Focused Recall	A method that combines the short-term recall used in 24HR with either a list of specific groups of fish to probe respondent to report the fish and amounts consumed the previous day. Mode may be in-person, telephone, or web. Automated instrument exists for collecting 24HR/food list combination. Advantages: Provides reasonably good estimates of consumption of fish for previous day. Single day used to describe average dietary intake of a group; multiple days (or when paired with FFQ data) used to estimate population distributions (NCI method described in Chapter 7). Relatively low respondent burden. Limitations: If using the 24HR/food list combination, it may be difficult to include universe of local and non-local fish on food list. Increased measurement error due to underreporting of foods consumed as part of mixed dishes. Requires respondent to estimate portion of fish consumed. Instrument: • The FCST (discussed in Chapter 4). A configurable computerized interview originally designed to record fish consumption information for American Indians/Alaska Natives that collects 24HR recall, and fish frequency estimates of annual consumption (including source and parts consumed), consumption by children, and consumption at gatherings and special events (Kissinger et al., 2010). The automated survey tool allows researchers to populate the software with fish species of interest and define the seasonality and forms of interest. The fish database includes 77 species of fish. Survey includes a booklet with images of a variety of species, portion amounts, and fish preparation methods. Available from EPA free of charge.
Food Record/Diary (FR)	A method in which the respondent records all the foods and amounts consumed as they are eaten, over 1-4 consecutive days. Although more than 4 days results in decreased reporting (Gersovitz et al., 1978), 7-day records are sometimes used. The interview modes may be mail or web. Portion size aids assist the respondent to report amounts collected or portions may be weighed by the respondent (weighed food record). Advantages: Food description and amounts are recorded in real time and food portions can be directly measured. Can be used to collect total diet or only fish consumption. Single day used to describe average dietary intake of a group; multiple days needed or combined with FFQ to estimate population distributions. Limitations: Respondent training required; subject to bias because respondent may change diet or reporting over time; requires respondent motivation and literacy; high respondent burden; increased measurement error with more days of reporting; requires dietary coding and processing to generate amount consumed; weighed food records require equipment and additional training. Instrument instructions will need to emphasize fish consumption. Instrument:

Method	Description
	• Fred Hutchinson Cancer Research Center (FHCRC). Multiple Day Food Record. Use of template requires a fee. Contact FHCRC Nutrition Assessment at (800) 460-7270 or (206) 667-4161 http://sharedresources.fhcrc.org/sites/default/files/MultipledayFoodRecord.pdf
Food Frequency Questionnaire (FFQ)	A method to assess usual frequency of consumption of each listed food over a reference period (e.g., months or a year). FFQs may be based on an extensive list of food items (total diet) or a short list of specific foods. Typically self-administered using paper- or web-based formats, but can be interviewer administered, in-person, or by telephone. May ask discrete portion sizes or assign default portions. Advantages: Estimate usual intake over long time period up to 1 year. Best suited to rank individuals according to usual consumption to assess association between dietary intake and disease risk. May be combined with 24HR or FR to describe usual intake. Forms can be optically scanned. Limitations: Many details of dietary intake not captured, portion sizes not accurate due to categorical nature. Significant measurement error. Food list must reflect universe of foods of interest (including mixed dishes) in order to capture usual intake. Complex to develop or customize to focus on fish consumption (McNutt et al., 2008). Relatively high level of literacy required. Instruments: • National Cancer Institute (NCI) Diet History Questionnaire (DHQ). Paper or web-based instrument in English and Spanish that consists of a list of 134 food items, including six questions about fish consumption. Reference period is past year or month. Freely available (http://riskfactor.cancer.gov/DHQ/webquest/index.html). • Harvard FFQ. Paper- or web-based instrument that consists of 140+ food items. Five questions ask about fish consumption Reference period is past year. Purchase and analysis cost is $15-$20 per questionnaire (https://regepi.bwh.harvard.edu/health/nutrition.html). • Block FFQ. Paper- or web-based instrument in English and Spanish that consists of a list of 127 food items. Five questions ask about fish consumption. Purchase and analysis cost on request (http://www.nutritionquest.com).
Fish Diet Screener	A short list of fish foods (usually 15-30 foods); may include portion size questions. Mode may be mail, in-person, phone, or web. Advantages: Low respondent burden. Should be used only to capture information on a group of specific foods or a single food component. Limitations: No estimation of upper percentiles of usual intake, substantial amount of measurement error, literacy required for mail or web administration. Instruments: • USDA/NHANES fish frequency questionnaire. Asks frequency of consumption of 29 specific fish. Reference period is past 30 days. Publicly available (http://www.cdc.gov/nchs/nhanes/nhanes_questionnaires.htm). • Seafood Assessment Survey. Asks frequency and amount of consumption of

Method	Description
	clams and mussels by season. (http://appliedresearch.cancer.gov/diet/shortreg/instruments/fialkowski_shellfish_assessment.pdf).

REFERENCES

Carroll, R. J., Midthune, D., Subar, A. F., Shumakovich, M., Freedman, L. S., Thompson, F. E., & Kipnis, V. (2012). Taking advantage of the strengths of 2 different dietary assessment instruments to improve intake estimates for nutritional epidemiology. *American Journal of Epidemiology*, 175(4):340-347.

Gersovitz, M., Madden, J.P., Smiciklas-Wright, H. (1978). Validity of the 24-hour dietary recall and seven-day record for group comparisons. *Journal of the American Dietetic Association,* 73, 48-55.

Kirkpatrick, S., Subar, A., Douglass, D., Zimmerman, T., Thompson, F., Kahle, L., George. S., Dodd, K., and Potischman, N. (2014) .Performance of the Automated Self-Administered 24-hour Recall relative to a measure of true intakes and to an interviewer-administered 24-h recall. *The American Journal of Clinical Nutrition.* doi: 10.3945/ajcn.114.083238.

McNutt, S., Zimmerman, T., Hull, S. (2008). Development of food composition databases for food frequency questionnaires (FFQ). *Journal of Food Composition and Analysis,* S20-S26.

Tooze, J.A., Midthune, D., Dodd, K.W., Freedman, L.S., Krebs-Smith, S.M., Subar, A.F., Guenther, P.M., Carroll, R.J., Kipnis, V. (2006). A new statistical method for estimating the usual intake of episodically consumed foods with application to their distribution. *Journal of the American Dietetic Association,* Oct;106(10):1575-87.

APPENDIX C. APPLICATION OF THE NCI METHOD MACROS

The following information has been extracted from the NCI Method SAS Macros and User Guides available at http://epi.grants.cancer.gov/diet/usualintakes/macros.html and described based on the NCI Method macro experience of the authors.

The macro variables to define for use of the MIXTRAN are:

- Required
  - Data= Specifies the data set to be used.
  - Response= Specifies the 24 hour recall variable (in our example, this is "GramsFish")
  - Foodtype= Specifies a name for the analysis, used to identify the output data sets. This value can be the same as the response variable
  - Subject= Specifies the variable that uniquely identifies each subject (in our example, this is "Participant")
  - Repeat= Specifies the variable that indexes repeated for each subject (in our example, this is "RecallDay")
  - Covars_prob= Specifies a list of covariates for the first part of the model that models the probability of consumption (in our example, these could include "Season," "Gender," "Age," "Bodyweight," "FishFreq")
  - Covars_amt= Specifies a list of covariates for the second part of the model that models the consumption-day amount (in our example, these could include "Season," "Gender," "Age," "Bodyweight," "FishFreq")
  - Outlib= Specifies a directory where output data sets are stored
  - Modeltype= Specifies the model. For the best estimates "corr" should be specified, though it may not be necessary or feasible. A statistician should be consulted to determine if an uncorrelated model would be appropriate. The possible values are:
    1. "null string" = fit correlated model
    2. "corr" = fit correlated model
    3. "nocorr" = fit uncorrelated model
    4. "amount" = fit amount-only model

- Optional
  - Seq= Specifies one or more sequence indicator variables to account for effects due to the sequence number of a subject's records. This cannot also appear in covars_prob or covars_amt (in our example this is "Seq")
  - Weekend= Specifies the weekend (Fri.-Sun.) indicator to account for a weekend effect. This variable can NOT also appear in covars_prob or covars_amt (in our example, this is "Weekend")
  - Vargroup= Specifies a variable that groups observations to allow the model to incorporate a separate residual variance parameter for each of these groups of

observations. If the output from this macro is to be used in the DISTRIB macro, then only the weekend variable can be used.

- Numvargroups= Specifies the number of groups defined by the vargroup variable. If the output from this macro is to be used in the DISTRIB macro and weekend is the "vargroup" variable, then the number of groups is 2.
- Replicate_var= Specifies the variable to be used in the replicate statement of PROC NLMIXED or the freq statement of PROC UNIVARIATE. The specified variable must be integer valued (this is akin to a survey weight)

Note that there are other macro variables that can be defined. Please see the User's Guide to determine their necessity for your analysis, although in most cases, we believe those listed above are the ones necessary for best estimates.

After the MIXTRAN macro runs, it produces a number of output data sets, two of which will then be used by the DISTRIB macro.

> "The DISTRIB macro uses results from the MIXTRAN macro and estimates the distribution of usual intake for episodically foods, foods consumed every day, and nutrients (Tooze et al., 2006). First, the DISTRIB macro reads data sets of parameter estimates predicted values output by the MIXTRAN macro. Then, Monte Carlo simulation of the random effect(s) is used to estimate the distribution of usual intake." (NCI Method, DISTRIB macro, version 2.1)

The macro variables to define for use of the DISTRIB macro are:

- Required
  - Seed= Specifies the seed for the random number used for the Monte Carlo simulation of the random effects u1 and u2
  - Nsim_mc= Specifies the number of repetitions to be used in the Monte Carlo simulation. For each subject, record will be output for each repetition.
  - Modeltype= Specifies the model that was used by the MIXTRAN macro to prepare the data for the DISTRIB macro. The value must be the same as the model declared for the MIXTRAN macro.
  - Pred= Specifies the name of the data set containing values calculated in the MIXTRAN macro. The data set will be named _pred_XX, where XX = Foodtype specified in MIXTRAN.
  - Param= Specifies the name of the data set containing the parameter estimates calculated in the MIXTRAN macro. The data set will be named _param_XX, where XX = Foodtype specified in MIXTRAN.
  - Outlib= Specifies the library reference to which the output data set of distributions will be written.
  - Subject= Specifies the variable that uniquely identifies each subject. Required when "weekend" is used in MIXTRAN. (In our example this is "Participant")
  - Food= Specifies a name for the analysis, used to label the output data set.

- NLoptions= Specifies a list of options to be added to all calls to PROC NLMIXED, for example: nloptions=qpoints=1 gconv=1e-12 itdetails. This may be necessary if the models fail to converge. For example, it may be necessary to raise or lower the quadrature points (qpoints) or the convergence criteria (gconv). Please consult the User's Guide and SAS Help for the NLMixed procedure for more information.

Note that there are other macro variables that can be defined. Please see the NCI Method macro documentation and user's guide available at http://epi.grants.cancer.gov/diet/usualintakes/macros_single.html to determine their necessity for your analysis, although in most cases, we believe those listed above are the ones necessary for best estimates.

The DISTRIB macro outputs a data set "_mcsim1" that can be used to estimate the distribution of UFCR for the total population and subpopulations that were included as covariates. If you are going to use this data file, you will need to comment out a statement in the DISTRIB macro where it deletes the data set. The macro contains a comment to show the user where this needs to be done. It is in the last procedure call of the macro.

REFERENCES

Tooze, J.A., Midthune, D., Dodd, K.W., Freedman, L.S., Krebs-Smith, S.M., Subar, A.F., Guenther, P.M., Carroll, R.J., Kipnis, V. (2006). A new statistical method for estimating the usual intake of episodically consumed foods with application to their distribution. *Journal of the American Dietetic Association,* Oct;106(10):1575-87.

APPENDIX D. FCST HARD-COPY QUESTIONNAIRE AND DOCUMENTATION

[HARD COPY QUESTIONNAIRE TO BE MODFIED FOR USE]

INTRODUCTION

We appreciate your willingness to participate in our fish and shellfish consumption survey. The information given in response to this questionnaire will help [*insert name of tribe/agency/organization conducting the study*] to understand the rates of fish and shellfish consumption and the species or types of fish and shellfish regularly consumed in [*insert area of interest*].

All of the information you provide to us is confidential. Your responses to the questions will be combined with those of others so that your answers cannot be identified. If you have any questions, you are welcome to call [*insert individual's name and affiliation*] at [*insert telephone number*].

It takes about [*insert time to complete*] to complete this questionnaire. There are four parts. In Part One, you will be asked how much fish and shellfish you ate yesterday. The second and third parts focus on particular types of fish and shellfish you eat, and where it was prepared. Finally, you will be asked for some general information.

Today's Date _____ _____ _____

Month Day Year

PART ONE: DIETARY INTAKE – 24 HOUR RECALL

Please tell us whether you ate any fish or shellfish yesterday. Think about what you ate for breakfast, lunch, dinner, and any snacks. Include any fish or shellfish that were part of sandwiches, salads, or other mixed foods (e.g., seafood pasta).

Q1. Did you eat any fish or shellfish yesterday from the time you woke up until the time you went to sleep for the night? Please include meals and snacks. Mark the box next to your answer.

- ☐ Yes
- ☐ No [Go to Part Two.]
- ☐ Don't Know [Go to Part Two.]

Q2. Did you eat more than one species or type of fish or shellfish yesterday? Mark the box next to your answer.

- ☐ Yes
- ☐ No
- ☐ Don't Know

[*Populate this table with species of interest to the study. Keep the other/unknown categories.*]

Q3. Please mark the box next to each species or type of fish or shellfish you ate yesterday.	**Q4.** Was the fish/shellfish prepared at home, or at a restaurant or another place? (Mark all that apply)	**Q5.** How much of each species did you eat yesterday? Write the total number of ounces (oz) or pounds (lbs).
☐ Salmon, canned	☐ Home ☐ Restaurant or Other Place ☐ Don't Know	_____ ounces (oz) pounds (lbs) [CIRCLE ounces or pounds]
☐ Salmon, fresh	☐ Home ☐ Restaurant or Other Place ☐ Don't Know	_____ ounces (oz) pounds (lbs) [CIRCLE ounces or pounds]
☐ Trout	☐ Home ☐ Restaurant or Other Place ☐ Don't Know	_____ ounces (oz) pounds (lbs) [CIRCLE ounces or pounds]
☐ Other finfish	☐ Home ☐ Restaurant or Other Place ☐ Don't Know	_____ ounces (oz) pounds (lbs) [CIRCLE ounces or pounds]
☐ Unknown finfish	☐ Home ☐ Restaurant or Other Place ☐ Don't Know	_____ ounces (oz) pounds (lbs) [CIRCLE ounces or pounds]
☐ Clams	☐ Home ☐ Restaurant or Other Place ☐ Don't Know	_____ ounces (oz) pounds (lbs) [CIRCLE ounces or pounds]
☐ Mussels	☐ Home ☐ Restaurant or Other Place ☐ Don't Know	_____ ounces (oz) pounds (lbs) [CIRCLE ounces or pounds]
☐ Shrimp	☐ Home ☐ Restaurant or Other Place ☐ Don't Know	_____ ounces (oz) pounds (lbs) [CIRCLE ounces or pounds]
☐ Other shellfish	☐ Home ☐ Restaurant or Other Place ☐ Don't Know	_____ ounces (oz) pounds (lbs) [CIRCLE ounces or pounds]
☐ Unknown shellfish	☐ Home ☐ Restaurant or Other Place ☐ Don't Know	_____ ounces (oz) pounds (lbs) [CIRCLE ounces or pounds]

PART TWO: FINFISH CONSUMPTION – SEASON, FREQUENCY, PORTIONS

For this part of the survey we will ask about your consumption of finfish and shellfish over the past 12 months. The first questions are about what species of finfish you ate, the amount you ate, and how often you ate each species over a year.

The amount of fish you eat and how often you eat it may depend on the time of year. For example, if there are seasonal differences in how often you eat fish, you may answer two different ways: when it is fresh and readily available, and when it is not in season. Or, if you believe there is no difference in how often you eat the fish, you can tell us how often you eat fish in general without regard to when it is in season. Please remember to include fish you eat at breakfast, lunch, dinner, and snacks. Do not include fish you eat at ceremonies or community gatherings and events, as we will ask you about those in a later section.

Q6. In the past 12 months, did you ever eat [*insert first species of interest*]?

- ☐ Yes
- ☐ No [Go to Q21.]
- ☐ Don't Know [Go to Q21.]

Q7. Is the amount of [*insert first species of interest*] you ate different when the fish was in season and is fresh and available compared to when the fish was not in season? Please keep in mind that the season is [*insert season for species of interest*].

- ☐ Yes
- ☐ No [Go to Q12.]

Q8. How often in the past 12 months did you eat [*insert first species of interest*] when the fish was **in season**?

_____ times per

(WRITE number of times on line)

- ☐ Day
- ☐ Week
- ☐ Month
- ☐ Season
- ☐ Year

(MARK box)

Q9. How much did you usually eat in a sitting when the fish is **in season**?

_____ ☐ Ounces (oz) ☐ Pounds (lbs)

(WRITE amount on line) (MARK box)

Q10. How often in the past 12 months did you eat [*insert first species of interest*] when the fish was **not in season**?

_____ times per

(WRITE number of times on line)

☐ Day
☐ Week
☐ Month
☐ Season
☐ Year

(MARK box)

Q11. How much did you usually eat in a sitting when the fish is **in not season**?

_____ ☐ Ounces (oz) ☐ Pounds (lbs) **[GO TO Q14.]**

(WRITE amount on line) (MARK box)

Q12. As your consumption did not vary by season, in the past 12 months, how often did you eat [*insert first species of interest*]?

_____ times per

(WRITE number of times on line)

☐ Day
☐ Week
☐ Month
☐ Season
☐ Year

(MARK box)

Q13. How much did you usually eat in a sitting?

_____ (WRITE amount on line)

☐ Ounces (oz)
☐ Pounds (lbs)

(MARK box)

Q14. What percent of the time did you eat [*insert species of interest*] with skin?

- ☐ Never
- ☐ 1 to 10% of the time
- ☐ 11 to 20% of the time
- ☐ 21 to 30% of the time
- ☐ 31 to 40% of the time
- ☐ 41 to 50% of the time
- ☐ 51 to 60% of the time
- ☐ 61 to 70% of the time
- ☐ 71 to 80% of the time
- ☐ 81 to 90% of the time
- ☐ 91 to 99% of the time
- ☐ All of the time (100%)
- ☐ Don't Know

Q15. What percent of the time did you eat [*insert species of interest*] eggs if they are available?

- ☐ Never
- ☐ 1 to 10% of the time
- ☐ 11 to 20% of the time
- ☐ 21 to 30% of the time
- ☐ 31 to 40% of the time
- ☐ 41 to 50% of the time
- ☐ 51 to 60% of the time
- ☐ 61 to 70% of the time
- ☐ 71 to 80% of the time
- ☐ 81 to 90% of the time
- ☐ 91 to 99% of the time
- ☐ All of the time (100%)
- ☐ Don't Know

Q16. What percent of the time did you eat the head, bones, or organs of [*insert species of interest*]?

- ☐ Never
- ☐ 1 to 10% of the time
- ☐ 11 to 20% of the time
- ☐ 21 to 30% of the time
- ☐ 31 to 40% of the time
- ☐ 41 to 50% of the time
- ☐ 51 to 60% of the time
- ☐ 61 to 70% of the time
- ☐ 71 to 80% of the time
- ☐ 81 to 90% of the time
- ☐ 91 to 99% of the time
- ☐ All of the time (100%)
- ☐ Don't Know

Q17. What percent of the time was the [*insert species of interest*] that you eat from the grocery store?

- ☐ Never
- ☐ 1 to 10% of the time
- ☐ 11 to 20% of the time
- ☐ 21 to 30% of the time
- ☐ 31 to 40% of the time
- ☐ 41 to 50% of the time
- ☐ 51 to 60% of the time
- ☐ 61 to 70% of the time
- ☐ 71 to 80% of the time
- ☐ 81 to 90% of the time
- ☐ 91 to 99% of the time
- ☐ All of the time (100%)
- ☐ Don't Know

Q18. What percent of the time was the [*insert species of interest*] that you eat caught by you, your family members, or friends?

- ☐ Never
- ☐ 1 to 10% of the time
- ☐ 11 to 20% of the time
- ☐ 21 to 30% of the time
- ☐ 31 to 40% of the time
- ☐ 41 to 50% of the time
- ☐ 51 to 60% of the time
- ☐ 61 to 70% of the time
- ☐ 71 to 80% of the time
- ☐ 81 to 90% of the time
- ☐ 91 to 99% of the time
- ☐ All of the time (100%)
- ☐ Don't Know

Q19. What percent of the time was the [*insert species of interest*] that you eat from restaurants?

- ☐ Never
- ☐ 1 to 10% of the time
- ☐ 11 to 20% of the time
- ☐ 21 to 30% of the time
- ☐ 31 to 40% of the time
- ☐ 41 to 50% of the time
- ☐ 51 to 60% of the time
- ☐ 61 to 70% of the time
- ☐ 71 to 80% of the time
- ☐ 81 to 90% of the time
- ☐ 91 to 99% of the time
- ☐ All of the time (100%)
- ☐ Don't Know

Q20. What percent of the time was the [*insert species of interest*] that you eat from other sources?

- ☐ Never
- ☐ 1 to 10% of the time
- ☐ 11 to 20% of the time
- ☐ 21 to 30% of the time
- ☐ 31 to 40% of the time
- ☐ 41 to 50% of the time
- ☐ 51 to 60% of the time
- ☐ 61 to 70% of the time
- ☐ 71 to 80% of the time
- ☐ 81 to 90% of the time
- ☐ 91 to 99% of the time
- ☐ All of the time (100%)
- ☐ Don't Know

[*Repeat questions 6-20 for each finfish species of interest and/or other finfish*]

PART THREE: SHELLFISH CONSUMPTION – FREQUENCY, PORTIONS

The following questions are about what species of shellfish you eat, the amount you eat, how often you eat each species, and what parts of the shellfish you consume.

Q21. In the past 12 months, did you ever eat [*insert first shellfish species of interest*]?

- ☐ Yes
- ☐ No [Go to Q35.]
- ☐ Don't Know [Go to Q35.]

Q22. Is the amount of [*insert first shellfish species of interest*] you ate different when the [*insert they were/it was*] in season and is fresh and available compared to when [*insert they were/it was*] not in season? Please keep in mind that the season is [*insert season for species of interest*].

☐ Yes
☐ No [Go to Q27.]

Q23. How often in the past 12 months did you eat [*insert first shellfish species of interest*] when [*insert they were/it was*] **in season**?

_____ times per
(WRITE number of times on line)

☐ Day
☐ Week
☐ Month
☐ Season
☐ Year

(MARK box)

Q24. How much did you usually eat in a sitting when [*insert they were/it was*] **in season**?

_____

☐ Ounces (oz)
☐ Pounds (lbs)

(WRITE amount on line) (MARK box)

Q25. How often in the past 12 months did you eat [*insert first shellfish species of interest*] when [*insert they were/it was*] not **in season**?

_____ times per
(WRITE number of times on line)

☐ Day
☐ Week
☐ Month
☐ Season
☐ Year

(MARK box)

Q26. How much did you usually eat in a sitting when [*insert they were/it was*] **in not season**?

_____

☐ Ounces (oz)
☐ Pounds (lbs)

[GO TO Q27.]

(WRITE amount on line) (MARK box)

Q27. As your consumption did not vary by season, in the past 12 months, how often did you eat [*insert first shellfish species of interest*]?

_____ times per

(WRITE number of times on line)

- ☐ Day
- ☐ Week
- ☐ Month
- ☐ Season
- ☐ Year

(MARK box)

Q28. How much did you usually eat in a sitting?

_____

- ☐ Ounces (oz)
- ☐ Pounds (lbs)

(WRITE amount on line) (MARK box)

Q29. What percent of the time did you eat the meat/whole body of the [*insert species of interest*]?

- ☐ Never
- ☐ 1 to 10% of the time
- ☐ 11 to 20% of the time
- ☐ 21 to 30% of the time
- ☐ 31 to 40% of the time
- ☐ 41 to 50% of the time
- ☐ 51 to 60% of the time
- ☐ 61 to 70% of the time
- ☐ 71 to 80% of the time
- ☐ 81 to 90% of the time
- ☐ 91 to 99% of the time
- ☐ All of the time (100%)
- ☐ Don't Know

Q30. What percent of the time do you eat other parts of the [*insert shellfish species of interest*]?

- ☐ Never
- ☐ 1 to 10% of the time
- ☐ 11 to 20% of the time
- ☐ 21 to 30% of the time
- ☐ 31 to 40% of the time
- ☐ 41 to 50% of the time
- ☐ 51 to 60% of the time
- ☐ 61 to 70% of the time
- ☐ 71 to 80% of the time
- ☐ 81 to 90% of the time
- ☐ 91 to 99% of the time
- ☐ All of the time (100%)
- ☐ Don't Know

Q31. What percent of the time was the [*insert shellfish species of interest*] that you ate from the grocery store?

- ☐ Never
- ☐ 1 to 10% of the time
- ☐ 11 to 20% of the time
- ☐ 21 to 30% of the time
- ☐ 31 to 40% of the time
- ☐ 41 to 50% of the time
- ☐ 51 to 60% of the time
- ☐ 61 to 70% of the time
- ☐ 71 to 80% of the time
- ☐ 81 to 90% of the time
- ☐ 91 to 99% of the time
- ☐ All of the time (100%)
- ☐ Don't Know

Q32. What percent of the time was the [*insert shellfish species of interest*] that you ate caught by you, your family members, or friends?

- ☐ Never
- ☐ 1 to 10% of the time
- ☐ 11 to 20% of the time
- ☐ 21 to 30% of the time
- ☐ 31 to 40% of the time
- ☐ 41 to 50% of the time
- ☐ 51 to 60% of the time
- ☐ 61 to 70% of the time
- ☐ 71 to 80% of the time
- ☐ 81 to 90% of the time
- ☐ 91 to 99% of the time
- ☐ All of the time (100%)
- ☐ Don't Know

Q33. What percent of the time was the [*insert shellfish species of interest*] that you ate from restaurants?

- ☐ Never
- ☐ 1 to 10% of the time
- ☐ 11 to 20% of the time
- ☐ 21 to 30% of the time
- ☐ 31 to 40% of the time
- ☐ 41 to 50% of the time
- ☐ 51 to 60% of the time
- ☐ 61 to 70% of the time
- ☐ 71 to 80% of the time
- ☐ 81 to 90% of the time
- ☐ 91 to 99% of the time
- ☐ All of the time (100%)
- ☐ Don't Know

Q34. What percent of the time was the [*insert shellfish species of interest*] that you ate from other sources?

- ☐ Never
- ☐ 1 to 10% of the time
- ☐ 11 to 20% of the time
- ☐ 21 to 30% of the time
- ☐ 31 to 40% of the time
- ☐ 41 to 50% of the time
- ☐ 51 to 60% of the time
- ☐ 61 to 70% of the time
- ☐ 71 to 80% of the time
- ☐ 81 to 90% of the time
- ☐ 91 to 99% of the time
- ☐ All of the time (100%)
- ☐ Don't Know

[*Repeat questions 21-34 for each species of interest and/or other shellfish*]

PART FOUR: GENERAL QUESTIONS

This is the final section of the questionnaire. It asks you for general information on fish and shellfish consumption followed by a few questions about you.

Q35. Sometimes people's consumption of fish and shellfish changes over time. Has the amount of fish and/or shellfish consumed by you or your family changed over the last 20 years?

- ☐ Yes
- ☐ No [Go to Q38.]
- ☐ Don't Know [Go to Q38.]

Q36. How has the amount of fish and/or shellfish you have eaten over the past 20 years changed?

- ☐ Eat more now
- ☐ Eat less now
- ☐ Eat different types now
- ☐ Prefer not to say [Go to Q38.]
- ☐ Don't Know [Go to Q38.]

Q37. Please tell us what you think has caused the change in the amount or type of fish or shellfish you now eat.

_____

_____

_____

_____

_____

Q38. In most years, how many ceremonies, large gatherings, or other community events do you attend where fish and shellfish is consumed? [*insert examples of events pertinent to the study population*]

Number of events per year _____ [If none or zero, go to Q41.]

Q39. At these gatherings, do you ever eat [*insert first species of interest*]?

- ☐ Yes
- ☐ No [Go to Q41.]
- ☐ Don't Know [Go to Q41.]

Q40. How much [*insert first species of interest*] do you usually eat in at these gatherings?

_____ ☐ Ounces (oz)
☐ Pounds (lbs)

(WRITE amount on line) (MARK box)

[*Repeat questions 39 and 40 for each species of interest and/or other fish/shellfish*]

Q41. What is your gender?

- ☐ Male
- ☐ Female

Q42. How old are you?

- ☐ 18 to 29 years
- ☐ 30 to 39 years
- ☐ 40 to 49 years
- ☐ 50 to 64 years
- ☐ 65 to 79 years
- ☐ 80 years or older

[*Change and insert age groups of interest to study*]

Q43. What is your height (in feet and inches) and weight (in pounds)?

Height: _____ feet _____inches

Weight: _____ pounds

Q44. What is your race? (Adjust categories to fit the target population)

- ☐ White
- ☐ Black or African American
- ☐ American Indian/Alaska Native
- ☐ Asian
- ☐ Native Hawaiian or other Pacific Islander
- ☐ Two or more races

Q45. Are you of Hispanic origin? (Adjust categories to fit the target population)

- ☐ No, not of Hispanic, Latino, or Spanish origin
- ☐ Yes, Mexican, Mexican American, Chicano
- ☐ Yes, Puerto Rican
- ☐ Yes, Cuban
- ☐ Yes, other Hispanic, Latino, or Spanish origin

Q46. What is your household annual income?

- ☐ $0-$10,000
- ☐ $10,001-$20,000
- ☐ $20,001-$30,000
- ☐ $30,001-$40,000
- ☐ $40,001-$50,000
- ☐ $50,001 and over

[*Change and insert income groups appropriate for study population*]

CONCLUSION

Thank you very much for your cooperation in participating in this survey. Your participation will significantly contribute to information needed to help protect our natural resources.

As mentioned in the introductory letter, we will be mailing you a brief follow-up survey [*insert time frame as decided on by the study*] that includes only Part One of this survey. [*If applicable for study - We will mail your check for $25.00 soon after we receive your completed follow-up survey.]*

Please call us if you have any questions.

Again, thank you very much.

Developing the Booklet of Photographs

Preparation of Species Photographs

Survey designers will determine the specific fish and shellfish species to be included in the survey. For analysis and exposure assessment, it is important to have a comprehensive list of fish species included in the survey. However, the corresponding respondent burden and the impact of assembling booklets with a large number of species photographs should be considered.

Sources

Once the species included are defined in the CAPI configuration, the Supervisor searches for images of specific fish species to include in the booklet. Various sources could be used to obtain photographs or images. Consultation with the local department of fisheries may be useful. Government fish and wildlife agencies often have catalogs or posters of local fish species. Local universities may also have this type of resource. Internet searches provide a wealth of sources for fish and shellfish images. Using readily available electronic images obtained from the web is cost-effective and efficient. However, images obtained from web searches should always be verified against another source. In addition, some websites have limited access to images. They may require subscriptions or the use of passwords. Also, the format and the quality of on-line images may vary greatly from one website to another. It may be preferable to use images from websites associated with academic, government, or commercial institutions. This would help ensure some degree of uniformity and authenticity in the images. Examples of some of these websites are provided below.

1. The Washington Department of Fish and Wildlife supplies images of Washington state fish and shellfish on the website. Unless otherwise indicated, the images are owned by the Washington Department of Fish and Wildlife or have been made available to the Department for public use. They may be used for non-profit or educational purposes provided that the Department or copyright holder is properly credited. Weblink: http://wdfw.wa.gov/fishing/salmon/identification.html

2. The U.S Food and Drug Administration's (FDA) Seafood Products Research Center and the Center for Food Safety & Applied Nutrition have developed the Regulatory Fish Encyclopedia. The Regulatory Fish Encyclopedia (RFE) is a compilation of data in several formats that assists with the accurate identification of fish species. Weblink: http://www.fda.gov/food/foodscienceresearch/rfe/default.htm

3. Other sources may include commercial fishery, hatchery or aquaculture websites.
   Weblink:http://www7.taosnet.com/platinum/data/light/species/species.html

Preparation of Seafood Portion Photographs

Models of seafood portions are important in determining the amount of each type of seafood consumed by respondents. In general, the survey attempts to quantify seafood consumption in terms of uncooked wet weights. Eliciting measurements of uncooked seafood consumption is problematic in that consumers often do not evaluate their consumption in terms of uncooked, but rather cooked seafood portions. Cooking can cause substantial volume changes, and using uncooked seafood model portions may cause interviewees to underestimate the amount of seafood they consume. Therefore, EPA decided that cooked seafood portion models should be shown to interviewees while seafood weights prior to cooking are to be used to quantify consumption.

Consequently, it is important that accurate weights of edible tissue (+/- 0.1 grams) be associated with each photograph. Again, the CAPI records consumption in terms of raw or uncooked weights. There is some variability associated with the tissue weights for model types (e.g. shellfish portions, canned clams and fish). Therefore, replicates of the tissue mass associated with a photograph (e.g. 6 clams, 6 cans of fish, etc.) should be weighed and the mean and standard deviation computed.

Types of Preparation Models

A variety of approaches are available for depicting seafood portions, these include life-like casts of seafood, rough models giving the interviewee a general idea of portion volume, or life size photos. Casts accurately depict portion size; however they can be difficult to prepare, particularly for cooked seafood. Further, preparation of multiple casts for multiple sets of interview materials is time consuming. Rough seafood volume models can also be useful, but uncertainty is created due to the discrepancy between the shape of the model and the actual shape of the seafood portion. Digital photographs do not communicate volume information to the same degree physical models do. However, photographs are easier to prepare than physical models, can accurately depict cooked seafood portions, can be readily organized, are easily replicated to provide multiple sets of interview materials, are easy to physically transport during the interview process, and can be readily transmitted via the Internet. For these reasons, EPA chose actual size photographs over other methods of portraying seafood portion size.

However, using online images for the photographic models is not desirable since the images will vary greatly from one another depending on the source. Moreover, it is difficult to determine if the portion size represented in the on-line image is approximately what an individual might consume, with larger or smaller portions being conceptually accessible multiples or fractions of the proposed portion. Also, it is not possible to associate an uncooked weight with the portion size represented in the on-line image.

Therefore, the guidelines presented below outline the best approach for the creation of photographic preparation and portion size models. The objective is to obtain raw

weights, prepare or cook the items, and photograph the cooked items in a standardized manner.

Preparation and Portion Size Models

The seafood to be photographed and the types of models to be created depend on the specific CAPI configuration. The supervisor will need to configure the CAPI software by designating a *ModelID*, Model Description, and a Portion Description for each photographic model. The *ModelID* provided will uniquely identify a physical portion model in the database. The Model Description and Portion Description provided should be helpful to the interviewer and respondent so as to select an appropriate physical model that best represents the amount consumed. The CAPI software can be configured so that multiple models can be assigned to a species.

Survey designers will need to determine how the fish and shellfish species are typically consumed. If a type of seafood can be prepared in multiple ways and portion sizes of these different preparations differ in weight, then it is important that models of each preparation type be created. For example, fish may be consumed as fillets, steaks, canned and in fish soups. Clams may be consumed fresh, canned, in fritters, and in clam chowders or soups. For shellfish in particular, the number of organisms depicted per model are selected so that interviewees could describe their consumption in terms of a reasonably small number of multiples of the model portion. For example, an individual consuming 24 manila clams at a sitting would do so by indicating that they consumed 4 model portions, the model portion being 6 clams.

Creating Models

A detailed plan should be prepared for each fish species and each type of preparation model that needs to be created. The plan should include details about the procurement of seafood items, different preparation methods, recipes, weights of representative seafood items, photo descriptions, etc. Organizing this information into a table would be helpful.

Seafood items should be purchased at local markets or ordered from local wholesale retailer. The items should be representative of what would be available to the target survey population. The raw materials needed for creating each model type and the preparation of the model are described below.

1. Top loading or triple beam balance: Used to record the weight of raw fish and shellfish tissue within a weight of +/- 0.1 g.
2. Kitchen facilities for preparing seafood
3. Exacto knives: Used to scrape edible tissue out of shells.
4. Chef's knife: Used to pare down fillet models.
5. Weighing trays or beakers: Used to hold tissue during the weighing process.

In addition to the items listed above, additional equipment needed may include pans, knives, steamers, or other utensils.

Obtaining weights

The raw seafood is weighed using a standard food scale and the weight recorded. Where an exact weight is not associated with a portion size (e.g. six shrimp vs. 8 oz fillet), the weight of six representative raw portions are measured to determine the mean standard deviation of the preparation. For soups and chowder preparations, it is important to note that the weights of interest are for the fish or clam tissue and not all other materials in the preparation (i.e. onions, potatoes, cream, etc.). The mean weight is entered into the database while individual portion weight is associated with the photographic model.

Seafood Preparation

If specific recipes or preparation instructions are needed, they should be obtained prior to the photo shoot. Again, the models of the various preparation types should be representative of what is available and consumed by Tribal members. A shopping list of additional ingredients (e.g. butter, lemon, breadcrumbs, flour, eggs, and oil) needed to prepare the seafood items according to the recipes or preparation is needed.

Taking Photographs

It is recommended that photos be taken in two photo sessions over 2-3 days. The number of photographic models and the resources available will determine the logistics for the photo shoot. The required items needed for the photo shoot are listed below:

1. Digital camera
2. Tape measure
3. Tripod
4. Neutral colored plates
5. White poster board
6. Color printer

Neutral colored plates (such as neutral blue) are used to stage the prepared seafood items for the photographs. White plates are not recommended since white seafood items might not be as visible. The same size plates and bowls are used consistently throughout the shoot. The plates are staged against a white poster board backdrop and on top of white poster board or tablecloth. To aid in judging the size of the portion it is helpful to have standard utensils in the picture as well.

The use of a digital camera is recommended since this allows for the immediate confirmation of photo quality photos, and eliminates the need for scanning traditional photographs. The digital camera is mounted on a tripod and remains in the same position throughout the photo shoot for consistency in angle, lighting, and ease of setup. If possible, the lighting is adjusted to eliminate most shadows and provide maximum illumination. A tape measure is included in photographs of seafood portions so that photographs can be scaled to actual size.

Once photographs of the preparation models are taken, the digital images are resized and color corrected using standard photo publishing software such as PhotoShop, Corel Paint, MS Paint, iPhoto, etc.) To resize the photos, the supervisor should set up (define) a crop that best defines the area of the largest plate and the tape measure with a little white space surrounding it. This crop area should be used to crop all images. Keeping the same crop dimensions will force the smaller servings to be proportioned correctly. Color correction is done as necessary. Most applications have an AUTOADJUST feature to simplify things. The images should be saved as medium resolution .jpg files (150dpi). The file images should be named with brief titles that relate to content (ex. Clam_raw, Clam_cooked). The images are then imported into a Microsoft Power Point presentation or other some other layout program. The photos can then be scaled to actual size in Power Point, by drawing a 1" line segment on the same slide as a photo, followed by scaling the photo until a 1" increment on the photographed tape measure was equal in length to the 1" line segment.

Booklet Assembly

To facilitate editing and rearranging of images, it is recommended that each fish species and each preparation model is formatted onto its own individual Power Point slide. All images are printed on its own page. Pages should be formatted for printing in a landscape format since this allows the most space for the image. Images should be printed on a laser color printer. Printing from an inkjet color printer will not produce sufficiently clear images. For durability, lamination of the pages is recommended. Depending on the available resources the images can be produced on glossy photo quality paper. However this increases the cost and overall weight of the product. Pages can be three-hole punched for assembly into a three ringed binder or bound in spiral notebooks.

Labeling and Arranging Photos

In formatting and assembling the booklet, it is important to consider the ease of administration and the facilitation of the flow of the CAPI interview. It is recommended that the images be arranged in the sort order designated by the supervisor during the CAPI configuration. Supervisors can use the *GroupID* and the *SpeciesID* in the CAPI configuration to assemble the Species Identification section of the booklet. They may choose to put captions with the Species Description defined in the CAPI or codes on the photographs such that they are easily identifiable and match up with the Library. Supervisors can use label the photographic models with the *ModelID* or the Model Description.

Since The CAPI interview has several sections that refer to same photographs throughout the interview, interviewers will have to flip back and forth between pictures. Therefore, it is necessary to prepare two booklets: one for species identification and one for the preparation models.

The Fish Consumption Survey Tool (FCST)
CAPI Survey
Description of the Supervisor Software

Overview of Survey Structure

The computer assisted personal interview (CAPI) software is based on a hard-copy consumption survey developed for the Suquamish Nation.

The CAPI asks questions about consumption of individual species of fish and shellfish as well as consumption of groups of fish and shellfish. Fish and shellfish are grouped on the basis of similarities in behavior and habitat (e.g. migratory species, marine bottom dwelling fish, freshwater predatory fish, saltwater shellfish). Questions that capture information at the group level minimize Respondent burden and avoid the collection of data with little variance and added value during analysis that will contribute to an assessment of health risks. For example, Respondents are asked to consider all types of salmon consumed when providing catch area locations. They are not asked to provide locations for each species of salmon consumed since salmon are grouped together due to similar behaviors and habitats and thus are most likely be found in similar areas that would share contaminant concentrations. At the end of different survey sections, an opportunity is given to record consumption of seafood species that were not included in any of the groups.

The survey consists of several section types:

1. A 24-hour recall section records a Respondent's seafood consumption during the day prior to the survey.

2. The seasonal consumption section records consumption of seafood over the course of a year. The seasonal consumption rate section of the survey proceeds by asking questions to establish yearly consumption rates for individual species within a group prepared in various forms. Questions on individual species are then followed with general summary questions about all of the species consumed within that group (e.g. Where are species in that group harvested from; What parts of organisms from the group are consumed?).

3. The children's consumption section allows male and female Respondents with children less than six years of age to record their children's seafood consumption. Information captured includes a child's age, weight, breast feeding history as well as seafood consumption patterns. Questions about children's finfish consumption are asked on the basis of seafood groups rather than individual species. Shellfish consumption is asked for individual species.

4. The general section of the survey records information on how a Respondent's seafood consumption has changed over time, consumption of seafood at special gatherings or celebrations, and general information such as approximate age, income, body weight, and height.

Introduction to the CAPI

The Seafood Consumption Survey is administered by computer assisted personal interview (CAPI) software developed in MS Access 2010. Most of the screen shots included in this Guide reflect the survey as configured by EPA for the Quinault Nation in March 2006.

The software has been designed so users can specify their own version of the CAPI. Versions may vary due to definition of seafood categories or groups, species included in a group, allowable harvest and catch locations, as well as text that further clarifies a question asked by an Interviewer to a Respondent. A user can have multiple interviewers conducting the same version of the CAPI in the field.

Each study will have one designated Supervisor's PC. This Guide will refer to the user or users of this PC as the Supervisor. The Supervisor Module allows a Supervisor to configure the CAPI software specifically for his/her study. The configured version of the CAPI is then deployed to a USB device (also known as a flash-drive or thumb-drive). The USB device is intended to be used for CAPI installation on Interviewer PCs. An Interviewer PC is referred to as a site in this Guide. The Supervisor Module also allows for the accumulation of data collected at multiple sites, once the interviewing phase is complete, into a single database on the Supervisor's PC. The functionality and use of the Supervisor Module is described in this Guide.

Computer Requirements

The designated Supervisor computer must have a *Pentium III or higher processor. It must have at least 128 MB of memory though 256 MB is preferable. The hard disk should have at least 200 MB available space*. The computer's operating system should be:
Windows 7. MS Access 2010 or higher should be installed on the computer prior to installing the Supervisor Module software.

Setup

To install the Supervisor Module software follow the steps below.

1) Download the Fish Consumption Survey Tool (FCST) zip file available on EPA's website (https://www.epa.gov/fish-tech/epa-guidance-developing-fish-advisories)

2) Once downloaded, click on **Supervisor Seafood.exe** and follow the steps to install the "Welcome to the Supervisor Seafood Setup Wizard."

The installation will copy the necessary software files to the local folder C:\SupervisorSeafood unless otherwise specified during installation. The files copied are: four Access 2010 databases of which two (*SupervisorSeafood_dat.accdb* and *SeafoodSurvey_dat.accdb*) contain the structure of the data collected during interviews and lookup tables (i.e. tables that contain possible answers to interview questions) and the other two (*SupervisorSeafood_app.accdb* and *SeafoodSurvey_app.mdb)* contain application code and objects such as queries, forms, and reports. Also copied is the file and Install.msi that will be used by the Supervisor Module for deploying to and accumulating data from the sites. A shortcut icon is copied to the Supervisor PC's desktop that will be used to initiate the module. Under the folder C:\ \Supervisor Seafood, the installation procedure will create four subfolders. The Export folder will contain a zipped copy of the *Seafood_dat.accdb* file for each deployed site. The Import folder will contain an unzipped copy of the *Seafood_dat.accdb* file for each imported site. The Translate folder will contain a copy of the *Seafood_dat.accdb* file for each site after the "Other Specified" seafood items have been translated to an item found in the library (more on this below). Items not asked about during an interview but reported by the Respondent must be translated before integrated with other site data. These databases in the Translate folder are integrated into the master database of all sites with completed interviews in preparation for analysis.

Once the installation program has completed the below files will be in the C:\ SupervisorSeafood folder.

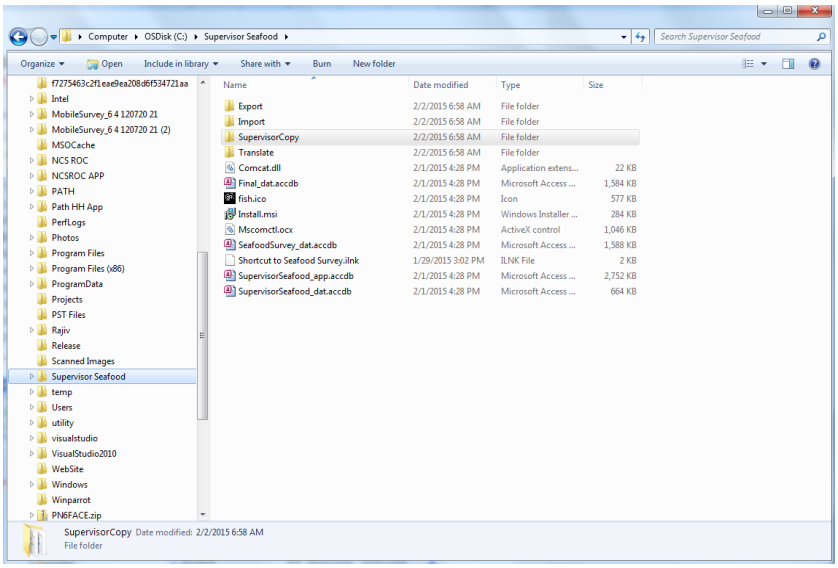


Figure 1. Folder structure after installation

The Supervisor Module

During installation an icon was copied to the Supervisor's PC desktop.

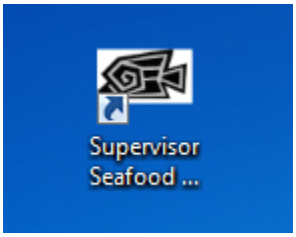


Figure 2. Desktop icon

Click on the icon to initiate the software. Click on "Yes" if prompted to allow the executable to install the software. When the Supervisor's Module is started for the first time, a prompt to "Enable content" will appear (Figure 3).

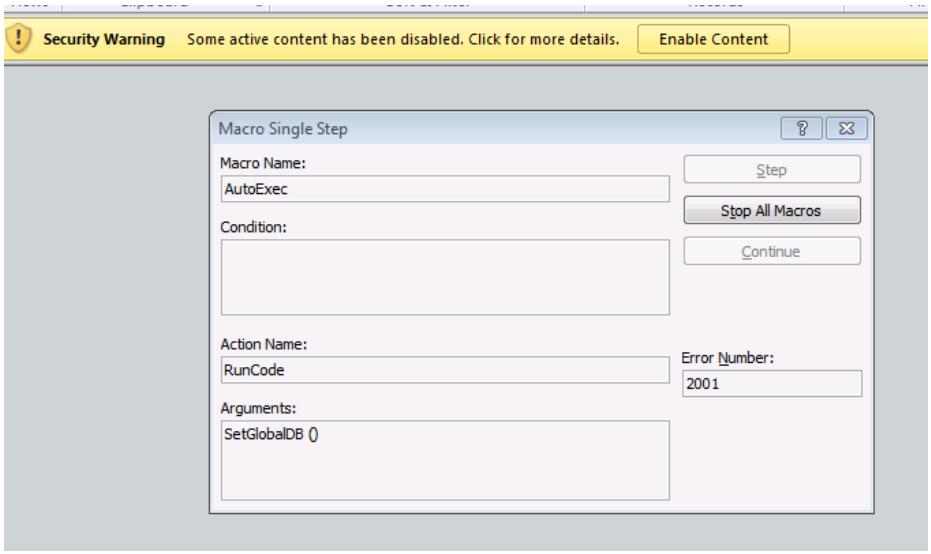


Figure 3: Enable Content prompt

Select the **Stop All Macros** button. Then select the **Enable Content** button in the yellow header. This step should not have to be followed after the first instance of starting the Supervisor software.

You will then be prompted for the Study number and name. The Study number allows for more than one study to utilize the same CAPI version. For example, a user may be planning on two studies, one of licensed anglers and one of the general population. Provide the name of the organization conducting the study (e.g., the State of Maryland, the Quinault Nation). Then provide the geographic region, area, or water body for which the study is interested in identifying harvest and catch areas (e.g. Puget Sound).

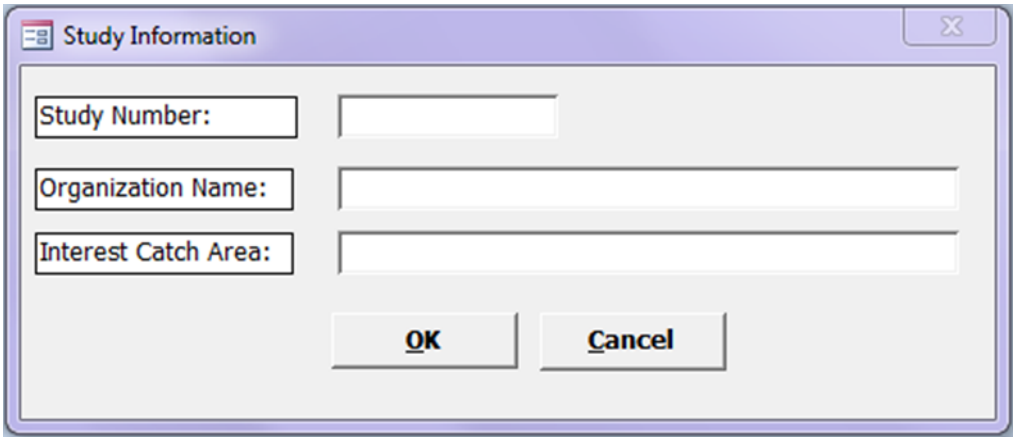


Figure 4. Study number and name prompt

Once this screen has been completed and the **OK** button pressed, a screen with a menu at the top will be displayed.

Figure 5. Supervisor Module's main menu

There are two tabs at the top : 1. File and 2. SCSM Menu. The file tab is the main functional menu for Access 2010. The working tab is the SCSM Menu tab.

The SCSM Menu has six items: ***File***, ***Library***, ***Configure***, ***Integrate***, ***Reports***, and ***About.*** ***Exit* is found under the *File* tab as well as under the *File* submenu.**

To exit the application, choose the ***Exit*** item on the ***File*** tab or the **File** submenu. The Windows exit control button  that appears in the top right hand corner of the screen has been disabled to prevent abnormal shutdown that might affect the integrity of the database.

Question Library

Select the ***Library*** menu item and then select Questions submenu to add or modify the library of questions. The library has been populated with records defined by EPA. Again, the screen shots of the library included in this Guide contain seafood used in the Quinault Tribal Consumption Survey.

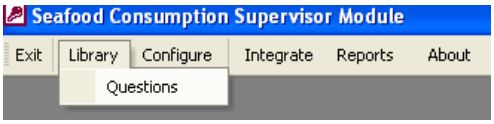


Figure 6. Library menu item

This menu item is only available if the CAPI configuration has **not** been finalized by the Supervisor (see page 27). The library is the master list of species that can be selected during configuration to be included in the CAPI. Additions or modifications should be done by an experienced Access programmer.

Select the ***Questions*** menu item to view the library.

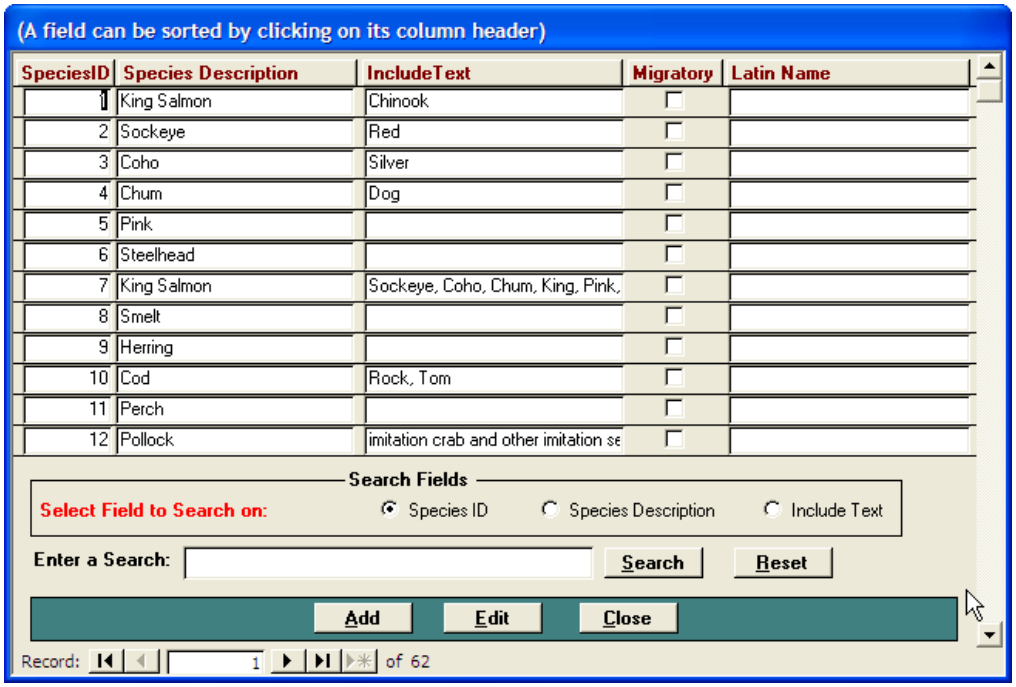


Figure 7. Species question library

Each record in the library has four fields associated with it that make the record unique:

1) *SpeciesID* is a unique numeric identifier assigned to each record. When adding a new record to the library, the system will calculate the next available number. This number, however, can be modified though it cannot duplicate an existing number. This ability to modify the *SpeciesID* allows a Supervisor to add a new record with a number (and presumably text) so that identical records can be maintained across studies.

Each CAPI consumption question or sets of questions are linked to a *SpeciesID* record selected by the Supervisor during configuration.

2) *SpeciesDescription* is the text that describes the species to be asked about in the CAPI. It is the text that is incorporated into the Interviewer Prompt in the CAPI. "King or Chinhook Salmon" is an example of a *SpeciesDescription*.

3) IncludeText is text that further explains the species to be asked about in the CAPI. The *SpeciesDescription* "King Salmon" may be asked about by itself or may be asked so that the Respondent considers specific types of salmon as well when answering the one question. This field might also be used to refer to common names of the *SpeciesDescription* that may be more recognizable to the survey population.

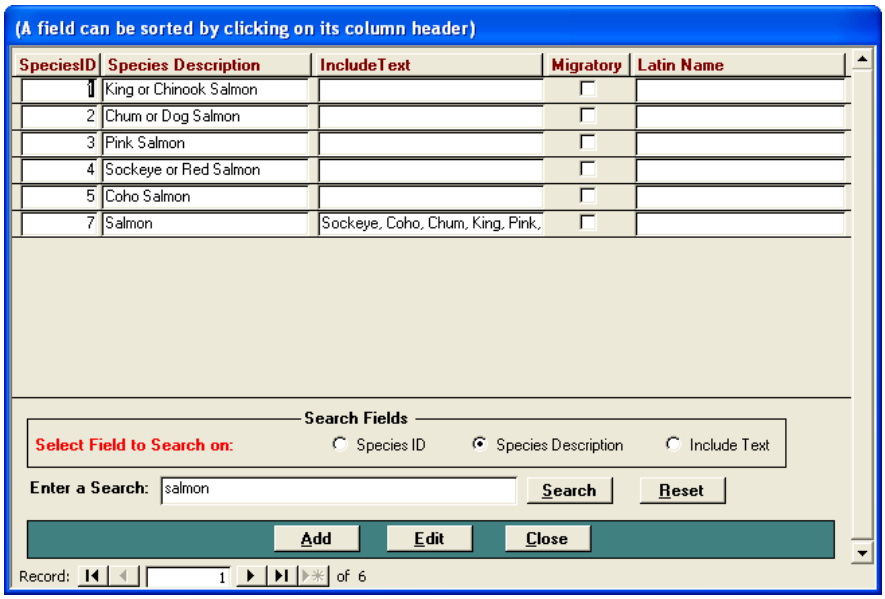


Figure 8. Salmon records in question library

While the columns are fixed length, the full text can be seen using the arrow keys when the field is selected.

Typically questions with identical *SpeciesDescription* values but different *IncludeText* values are selected during configuration to be included in different seafood frequency sections of the CAPI. The Salmon record, *SpeciesID* "1", if chosen by the Supervisor during configuration to be asked the 24-Hour Recall section will appear as in **Figure 9.**

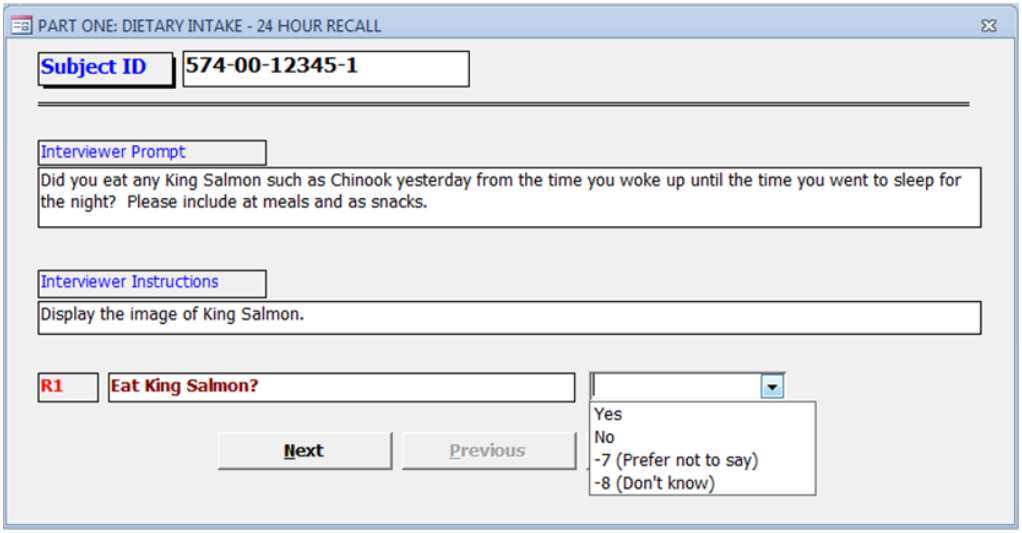


Figure 9. *SpeciesID* "1" was chosen to be included in the Recall section

The Salmon record, *SpeciesID* "1", if chosen by the Supervisor during configuration to be asked in the seasonal rates section will appear as:

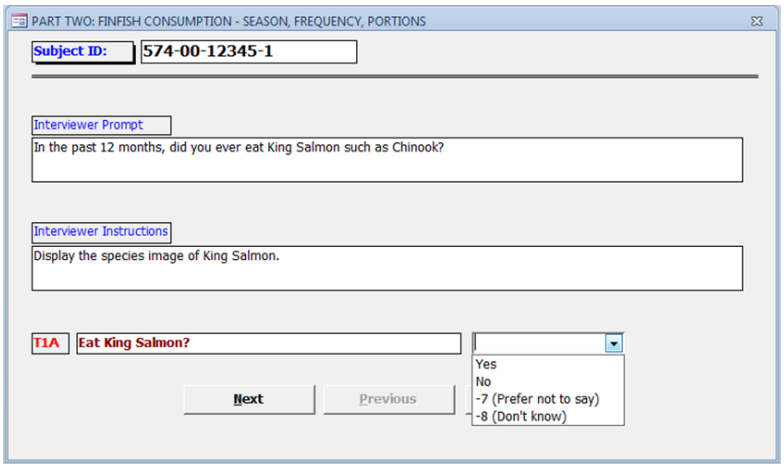


Figure 10. *SpeciesID* "1" was chosen to be included in the Rates section

4) *Migratory* specifies whether a species is considered migratory or resident of the region. The *SpeciesDescription* and *IncludeText* could be identical across two records but one species could be migratory and the second one not. This allows for a species, such as "Salmon," that would be considered migratory to one geographic region but in a different location might be lake locked and unable to migrate. The differentiation between records allows for more specific statistical analyses leading to better estimates.

As explained above when the ***Library*** sub-menu item is chosen a library form window ***Questions*** will be displayed. As is indicated in the title of the library window, you can sort one column in ascending order by clicking on the column header of a text field. The default sort order is by *SpeciesID*. Any one of the first three columns can be searched by completing the bottom third of the form. To find questions that reference the species "salmon", for example, choose the *SpeciesDescription* field and type "salmon" in the text field below. The search is not sensitive to case.

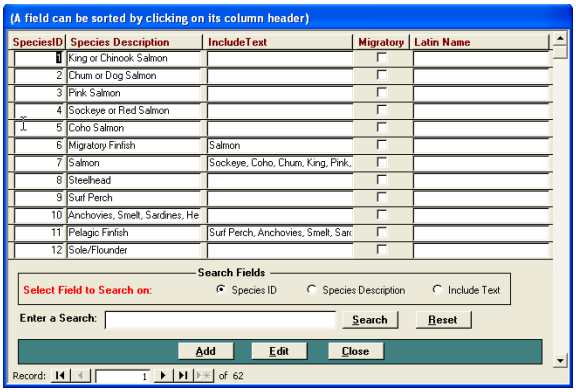


Figure 11. Searching for Salmon questions

Press the **Search** button to start the search. Resulting records containing the word "salmon" will be displayed.

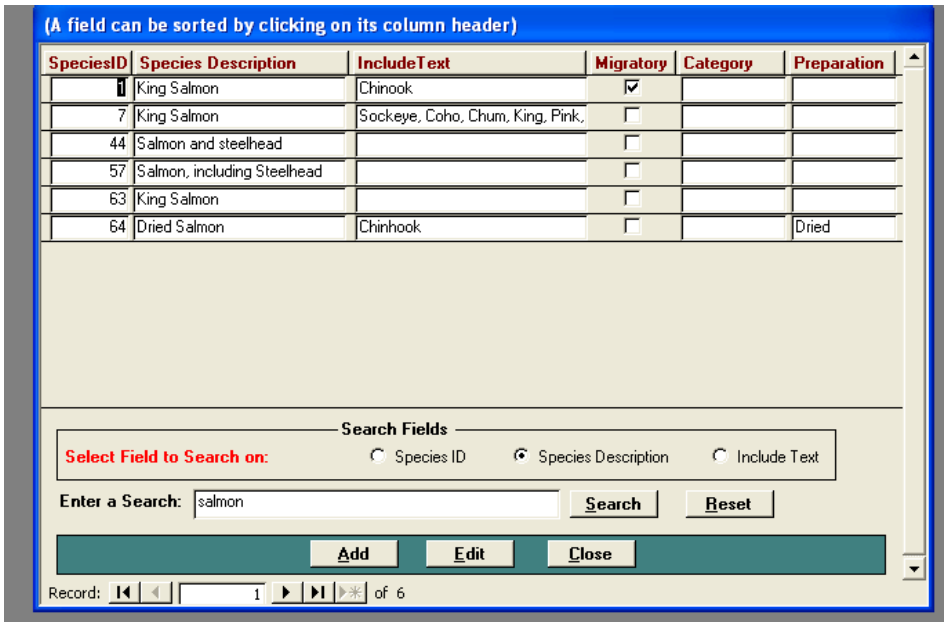


Figure 12. Search results for "salmon"

To display all the records again or to begin a new search, press the **Reset** button.

If the Supervisor has not published a final version of the CAPI, he/she may add a new record to the Species Question Library. To add a record, press the **Add** button to bring up the Add Species Question window.

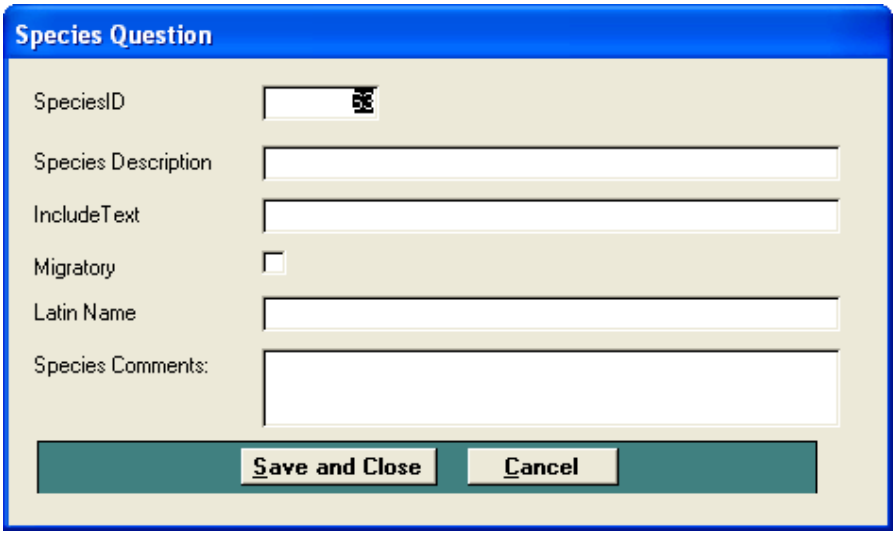


Figure 13. Adding a new species question to the library

The system will assign the next *SpeciesID* number though this can be modified to a number that does not already exist in the library. An error message will appear if you try to save a record with a *SpeciesID* that already exists. Press the **Save and Close** button to save the new record or press the **Cancel** button to not save the record and return to the library form.

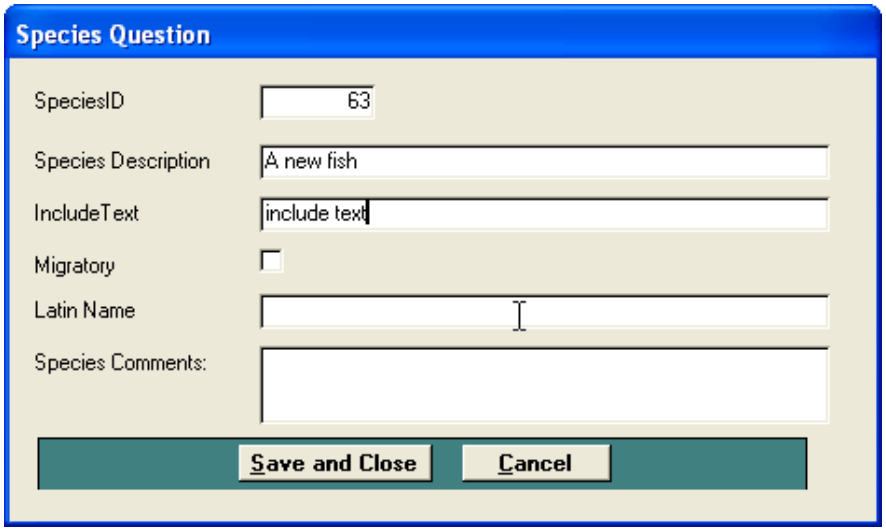


Figure 14. Adding a new record to the Species Question Library – a completed form

If the Supervisor has not published a final version of the CAPI, he/she may edit an existing record to modify the *Species Description*, *IncludeText*, *Migratory*, *Category*, and *Preparation* fields. Use the mouse to highlight a record. Press the **Edit** button to open up the edit window.

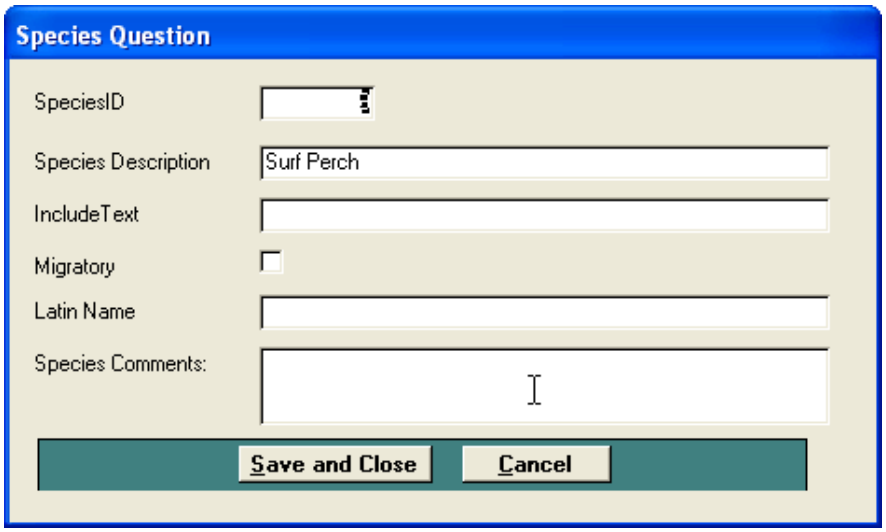


Figure 15. Editing an existing library record

Press the **Save and Close** button to save the edited record or press the **Cancel** button to not save the edit and return to the library form.

On the library form, press the **Close** button to return to the main menu.

Configuration

Select the ***Configure*** menu item and the ***Datasheets*** submenu item to configure your version of the CAPI.

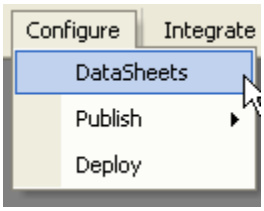


Figure16. Configure menu

To configure the software to your specific survey specifications, you will be taken through seven screens to customize the survey species questions. Use the **Next** and **Back** navigation buttons at the bottom of each of screen to move to the next screen or move back to the previous screen.

You will be asked to modify seven data tables that will define and control the questions that will be asked during the CAPI. For each of the tables you will be presented with a datasheet that is made up of rows and columns. Similar to a standard spreadsheet, each column represents a field of one type of information and each row represents an entry. Use the mouse or tab and arrow keys to navigate within the datasheet both vertically and horizontally. The right-facing arrowhead in the first column will indicate the row on which you are currently focused on. To delete a row, position the mouse in the left most column of the row (the whole row should be highlighted) and press the delete key. Sometimes, the text may appear to be truncated. This is because of the column width on the form. To increase the column width, position the mouse on the right box of the column header and drag it to the right to increase the display area. Note, however, increasing a column width will affect how many fields can appear on a page.

Other relevant navigation controls and indicators are indicated on the screen below.

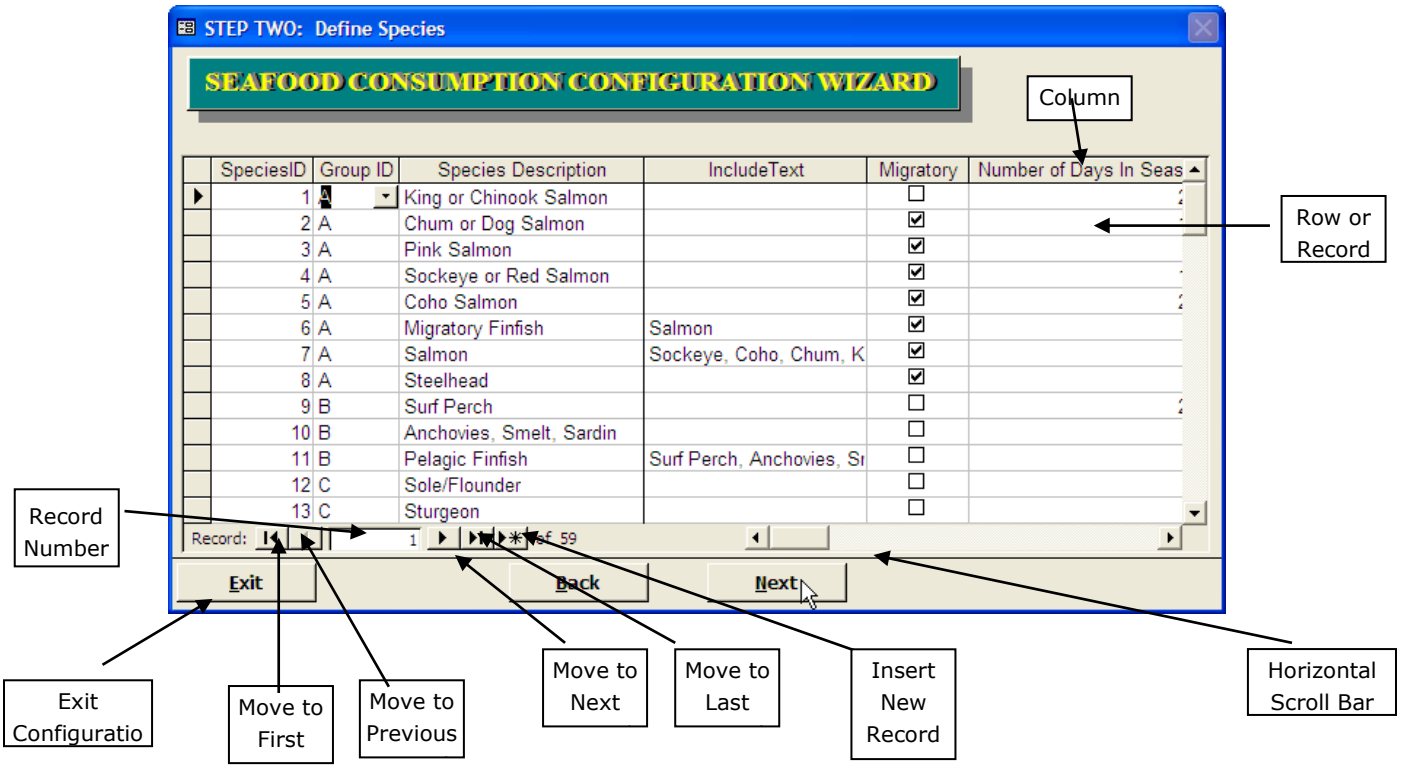

STEP TWO: Define Species
SEAFOOD CONSUMPTION CONFIGURATION WIZARD
Column
Row or Record
SpeciesID
Group ID
Species Description
IncludeText
Migratory
Number of Days In Seas
1 A King or Chinook Salmon
2 A Chum or Dog Salmon
3 A Pink Salmon
4 A Sockeye or Red Salmon
5 A Coho Salmon
6 A Migratory Finfish Salmon
7 A Salmon Sockeye, Coho, Chum, K
8 A Steelhead
9 B Surf Perch
10 B Anchovies, Smelt, Sardin
11 B Pelagic Finfish Surf Perch, Anchovies, Sr
12 C Sole/Flounder
13 C Sturgeon
Record:
1
of 59
Exit
Back
Next
Record Number
Exit Configuratio
Move to First
Move to Previous
Move to Next
Move to Last
Insert New Record
Horizontal Scroll Bar

Figure 17. Datasheet layout

The first datasheet screen (**Figure 18**) asks you to specify the groups for the fish and shellfish sections. Group assignment provides a means to organize fish according to similar characteristics. The group identifier is used by the CAPI to determine the order in which questions are asked within a section. (See the "Overview of Survey Structure" for further discussion of groups and how they are used.)

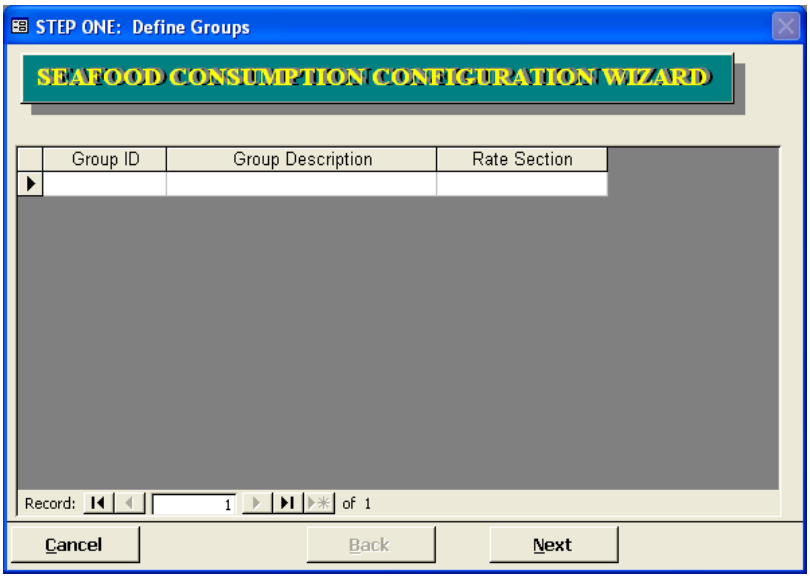


Figure 18. Group description

Groups should contain similar seafood items

Group Description is limited to 20 characters. Under the *Rate Section* column, choose which Rate Section of the Survey (Finfish, Shellfish, Other) the species in this groups should be asked about.

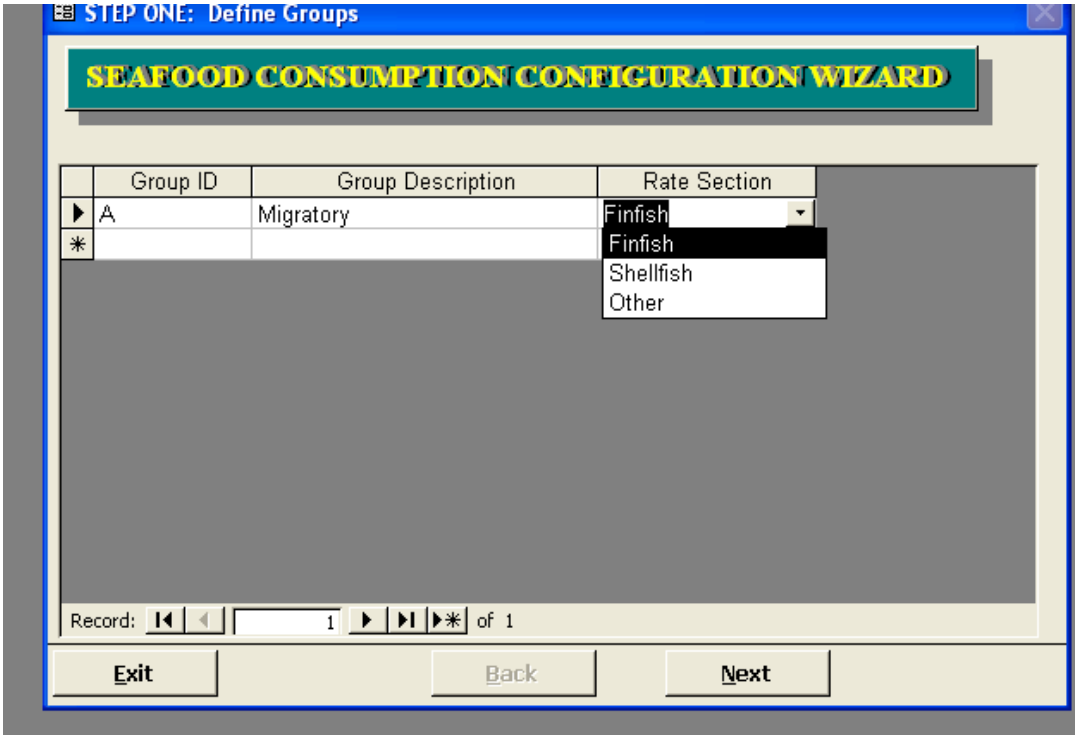


Figure 19. Rate Section designation

Respondents are asked to provide seasonal rates for Finfish but not for Shellfish. The body parts eaten are also asked differently for Finfish and Shellfish; Finfish parts are asked at the group level while Shellfish parts are asked at the species level. Finfish parts are pre-defined in the CAPI while Shellfish parts to be asked about are specified during this configuration process in a later datasheet. Finfish and Shellfish species types should always be mutually exclusive. For the example survey, the following groups were defined.

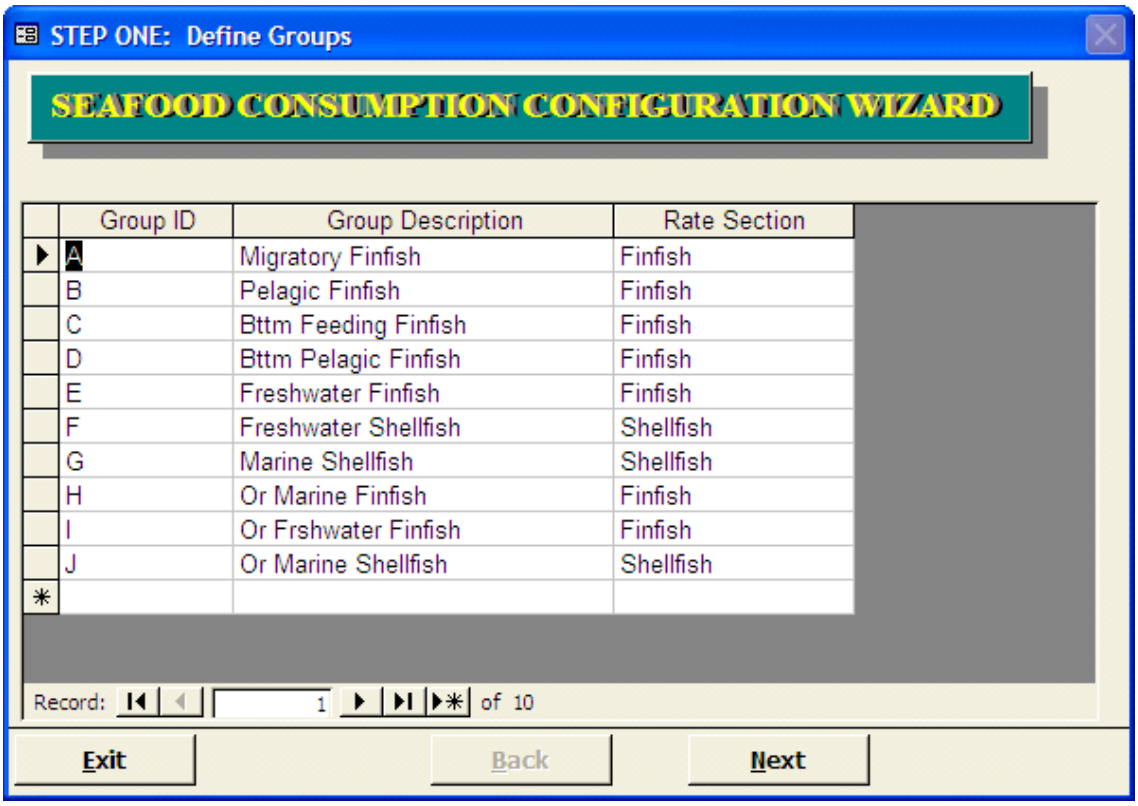


Figure 20. Definition of groups in Quinault Survey

Again, the *Group ID* determines the order in which species belonging to a group are asked in the CAPI. Within the *Rate Section* specified groups are asked in alphabetical order. In **Figure 20** species questions belonging to Group "A" will be asked before "B" which will be asked before "C" and so on. In the "Other" Rate Section Group "F" questions will precede Group "G" questions.

Press the **Next** button to advance to the next datasheet screen.

The second datasheet screen (**Figure 21**) will ask you to define which species will be asked in the four different sections of the CAPI (i.e. 24-hour recall, seasonal finfish and annual shellfish, child, and gatherings or special events) that ask about consumption.

In the *SpeciesDescription* column press the downward facing arrowhead to the right of the column to display records from the question library. They are sorted alphabetically. The drop down box displays both *SpeciesDescription* and *IncludeText* fields.

Figure 21. Species defined in the Question Library appear in drop down box

You can type ahead to find the species. For example to find "Steelhead", typing "st" will bring you to "Steelhead" line. To search the *SpeciesDescription* and *IncludeText* by keywords, press the **F5** key. A screen similar to the Question Library search screen (**Figure 12**) will be opened. Use the search functionality explained in the Question Library section, select the record containing the desired species, and press **OK** to return to the second configuration datasheet, "Define Species". The selected species will be filled in. Or press Cancel from the Question Library form to return to the second datasheet without selecting an item.

Note, any changes made to the corresponding record in the Question Library will not be reflected in a *SpeciesID* already selected on this datasheet. For example, if after you select *SpeciesId* "60", as in **Figure 22**, you decide you want to designate that fish as migratory. You must edit the record for *SpeciesId* "1" in the Question Library, come back to the second datasheet, delete the old *SpeciesId* "1" record, and add the modified *SpeciesId* "1" to the datasheet to pick up the new value of migratory.

Once you select the species, the fields *SpeciesID*, *Species Description*, *IncludeText*, and *Migratory* are filled in by the system. These cannot be modified. Next, select the appropriate group in the *Group ID* column by clicking your mouse on the downward facing arrowhead to the right of the column.

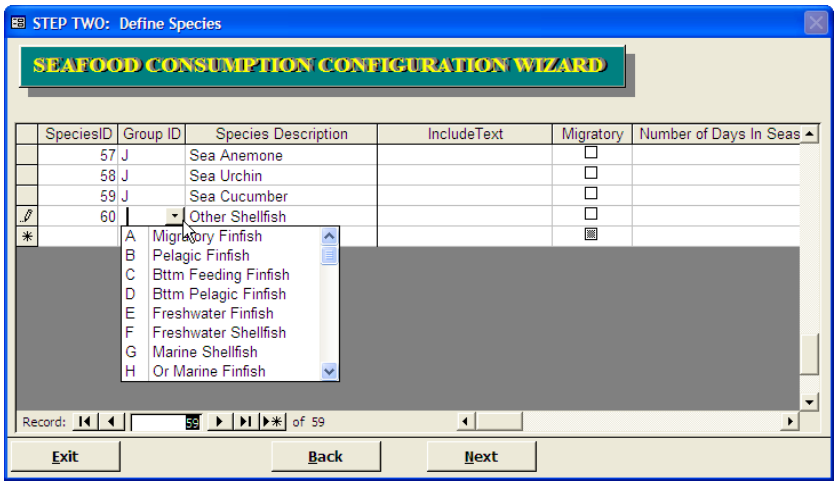


Figure 22. Group the new species question

As previously mentioned, Respondents are given an opportunity to report consumption of finfish and shellfish differently for when an item is in season versus out of season. Two values control the question flow, the Interviewer prompt and ultimately the calculation of consumption rates: *Number of Days in Season* and *In Season Text*.

Certain migratory species may return to their spawning areas at multiple time points over the course of a year. The Supervisor should consult with local fisheries biologists or other knowledgeable individuals to determine the commercial importance of these different migrations. If the commercial importance of one migration period vastly outweighs others, then the duration of that migration should be used to determine the "in season" time period. For example, Puget Sound Chinook salmon have spring and summer/fall migrations, however the commercial importance of the summer/fall migration greatly outweighs that of the spring migration. Consequently, the duration of the summer/fall migration should be used to determine "in season" consumption rates. Use the value zero "0" in the "Number of Days in Season" column to denote that there is no season for a given *SpeciesID*.

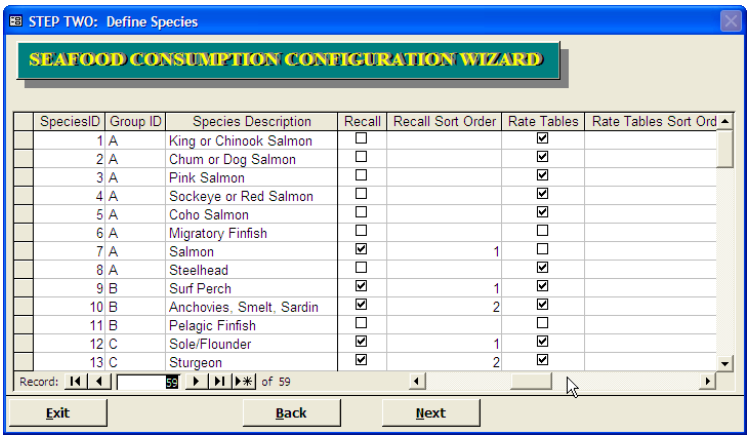


Figure 23. Providing season information

For each row or record in this datasheet, there are four checkboxes for each section in the survey that represent the four sections of the survey where consumption is ascertained: 24-hour Recall, seasonal and annual rates, child, and gatherings/special events. Each section checkbox has a corresponding sort order field. Use the horizontal scroll bar to move to the right to view these fields. Note the *SpeciesID*, *GroupID*, and *Species*

Description fields are frozen.

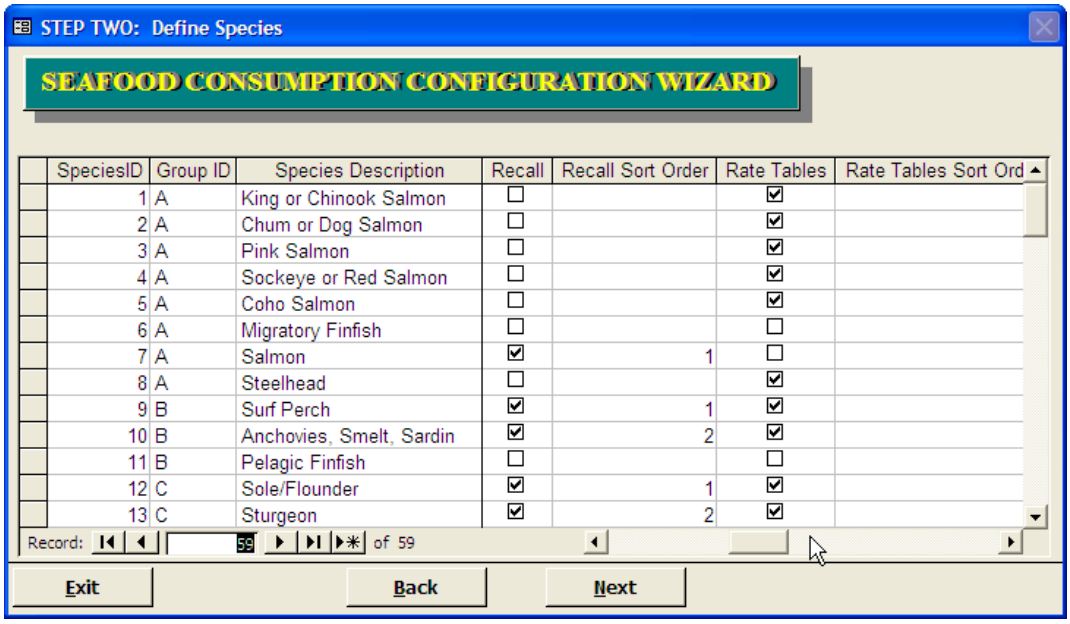


Figure 24. Horizontally scrolling to the right reveals Recall fields

Use your mouse key or tab to move to the box and hit the space bar so a checkmark appears in the checkbox. The checkmark will indicate that the item will be asked in the corresponding section of the survey. After each section's corresponding checkbox there is a field where you specify a number that will determine the order in which an item will be asked in each section of the survey within the *GroupID* field.

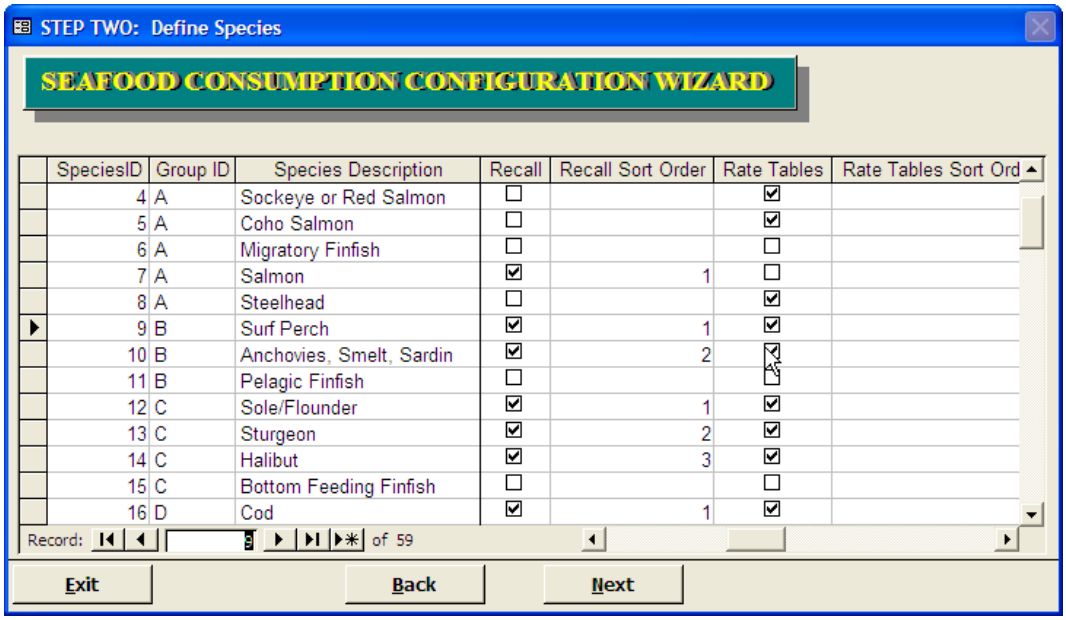


Figure 25. Selected questions to be asked in Recall Section

For example, in the Quinault Survey, Halibut, a bottom-feeding fish (Group "C"), is asked about in the 24-hour Recall (**Figure 25**). One food item entry or row in the datasheet will denote this. By checking off the *Recall* checkbox, the CAPI software will ask about eating Halibut in the section. A value of 3 in the *Recall Sort Order* column directs the software to ask about eating Halibut in the past 24-hours after Sturgeon, the second Group "C" fish asked about. Sole/Flounder is the first fish in Group "C" asked about in the 24-hour Recall.

If you scroll to the right, the first three columns again are frozen (**Figure 26**). Now, you can see Sole/Flounder is first Group "C" fish asked in the rate consumption table, Sturgeon, the second and Halibut the third.

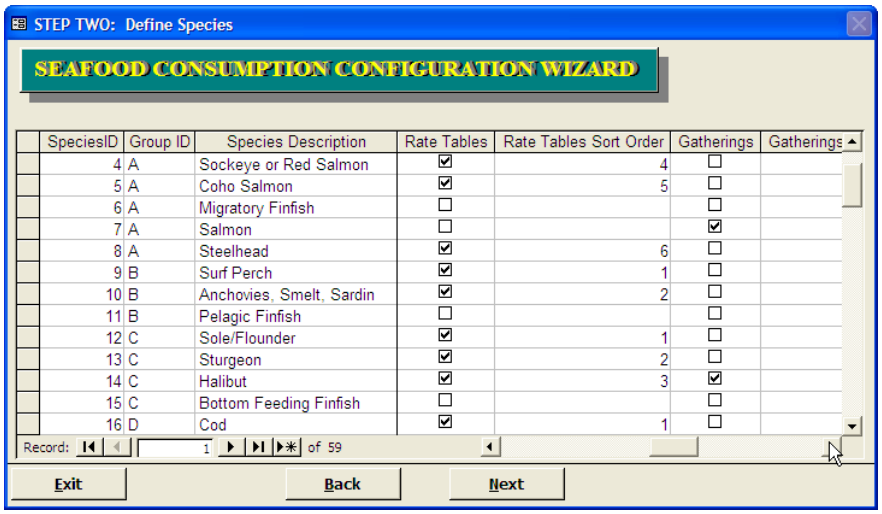
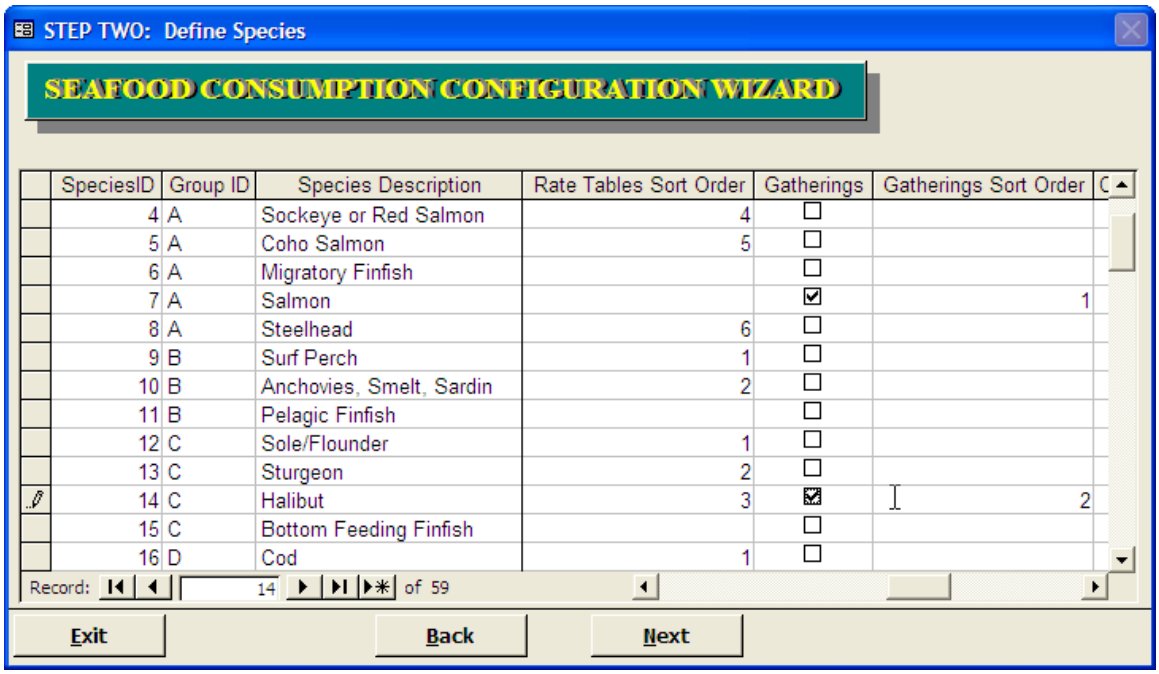


Figure 26. Halibut will be the third bottom-feeding fish asked in Consumption Rate section

Scrolling horizontally reveals that consumption at special events and gatherings of only Halibut, not Sole/Flounder and Sturgeon, will be asked about in the last section of the CAPI since the box it is the only one checked off in the Gatherings column.



SpeciesID	Group ID	Species Description	Rate Tables Sort Order	Gatherings	Gatherings Sort Order
4	A	Sockeye or Red Salmon	4	☐	
5	A	Coho Salmon	5	☐	
6	A	Migratory Finfish		☐	
7	A	Salmon		☑	1
8	A	Steelhead	6	☐	
9	B	Surf Perch	1	☐	
10	B	Anchovies, Smelt, Sardin	2	☐	
11	B	Pelagic Finfish		☐	
12	C	Sole/Flounder	1	☐	
13	C	Sturgeon	2	☐	
14	C	Halibut	3	☑	2
15	C	Bottom Feeding Finfish		☐	
16	D	Cod	1	☐	

Figure 27. Halibut asked about during the Gathering and Special Events section

The four sections of the survey that ask about consumption address consumption of species in different ways. Looking at the **Figure 28** below at salmon for example, you can see the 24-hour recall portion of the survey ascertains consumption of similar species as a collection, while the seasonal consumption portion of the survey generally ascertains consumption at the individual species level. In the 24-hour recall portion of the survey, collective salmon consumption as indicated by *SpeciesID* "7" is asked about. See **Figure 9** to view the resulting Interviewer prompt. While in the seasonal consumption portion of the survey, salmon consumption is recorded on a species specific basis. There will be six sets of seasonal consumption questions, one for *SpeciesID*s "1" - "6".

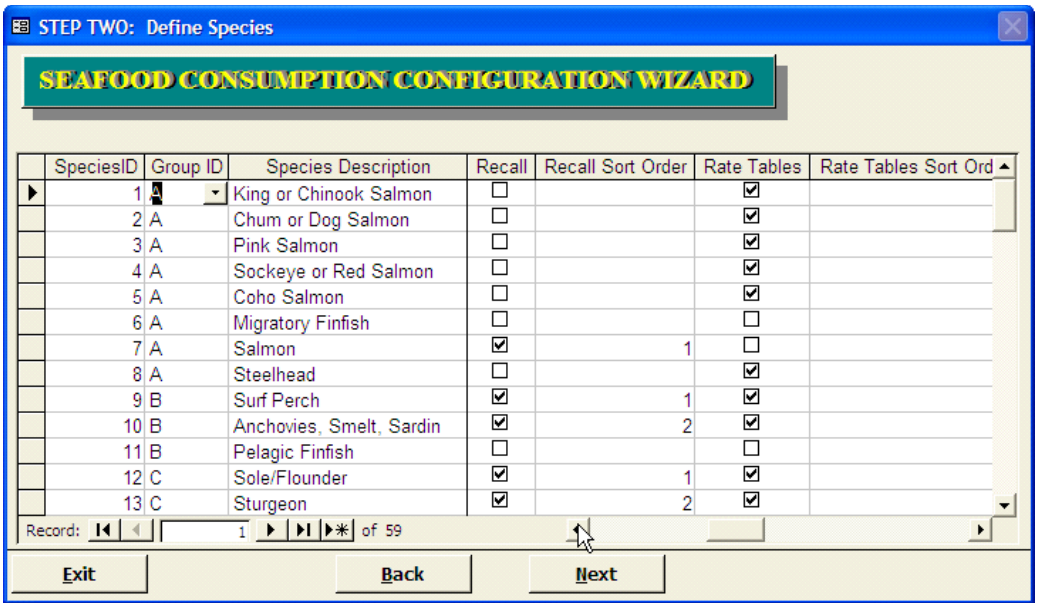


Figure 28. Different Salmon *SpeciesIDs*

Please note that finfish is asked at the group level in the child section and therefore should be configured appropriately in this datasheet. This design matches the Suquamish Nation survey of which this CAPI automated. To represent this in the table, enter one food item row for each collective group description to be asked as shown in the last five rows below. Only Groups designated to be asked in the Finfish or Shellfish Rate Sections can be asked in the child section of the CAPI.

Figure 29. Selected finfish items asked in child section

Shellfish items asked in the Child Section are identified in **Figure 30**.

Figure 30. Shellfish asked in child section

Press the **Next** button to proceed to the next configuration screen.

On the third datasheet screen (**Figure 31**) you itemize the photographs that will be used when conducting the interview to capture portion serving size of the item prepared a certain way. Photographs will be included in a booklet for the interview. Each photograph represents a standard portion size of a seafood item presented in a certain form such as a filet, as part of a soup, or fritters. The Respondent reports the amount he or she normally eats relative to the serving size represented in the picture (e.g. half the displayed amount, the displayed amount, twice the displayed amount). Each survey needs to develop a set of photographs for use in their survey. Guidance for developing the booklet is included with the FCST. Examples are included in the supplemental Microsoft PowerPoint file "FCST Example Species and Portion Size Photographs.ppt."

Enter one row for each unique photograph to be used during an interview. A serving size photograph can be used for multiple species. A species may be associated multiple photographs as defined on the next configuration datasheet.

The *Photo ID* provided will uniquely identify a photograph of a portion of a seafood item prepared in a certain form in the database. The *Preparation Form Type* and *Portion Description* provided should be helpful to the Interviewer and Respondent so as to select an appropriate photograph that best represents the amount and form consumed. The *Raw Gram Wt* will be used in the statistical reports to describe consumption amounts. Note the *Cooked Weight* is reserved for future use.

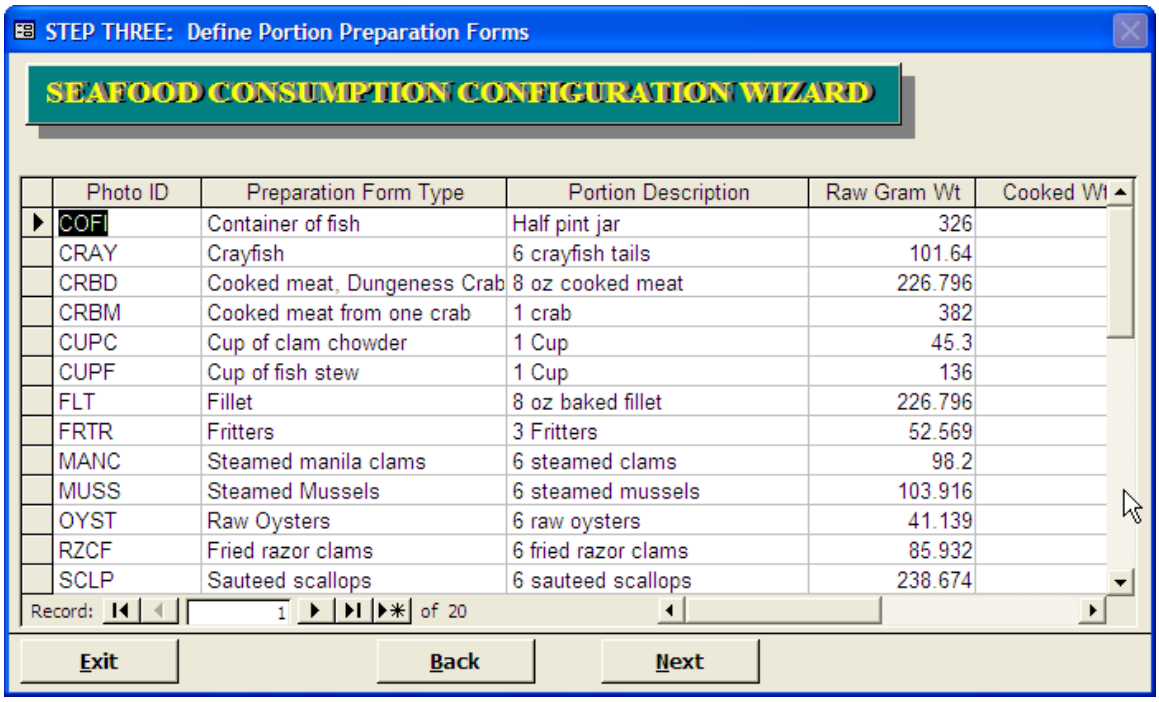


Figure 31. Defined portion models for Quinault CAPI

Press the **Next** button to proceed to the next configuration screen.

The fourth datasheet screen (**Figure 32**) is where you select which of the photographs defined in the previous datasheet screen are valid for the species questions included in the CAPI. In the first column, select the species for which you would like to associate a photograph. The drop down box will display the *SpeciesDescription* and *SpeciesID*.

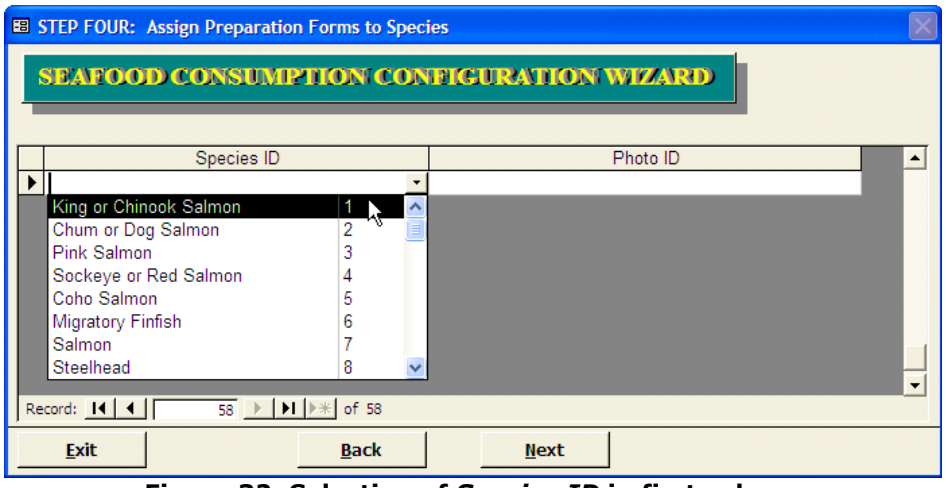


Figure 32. Selection of *Species ID* in first column

In the second column, select the portion photograph to be assigned to each species. You can assign multiple photographs to a species; each *Species ID/Photo ID* combination will be on its own row. *Preparation Form Type, Photo ID* and *Portion Description* are displayed in the drop down box.

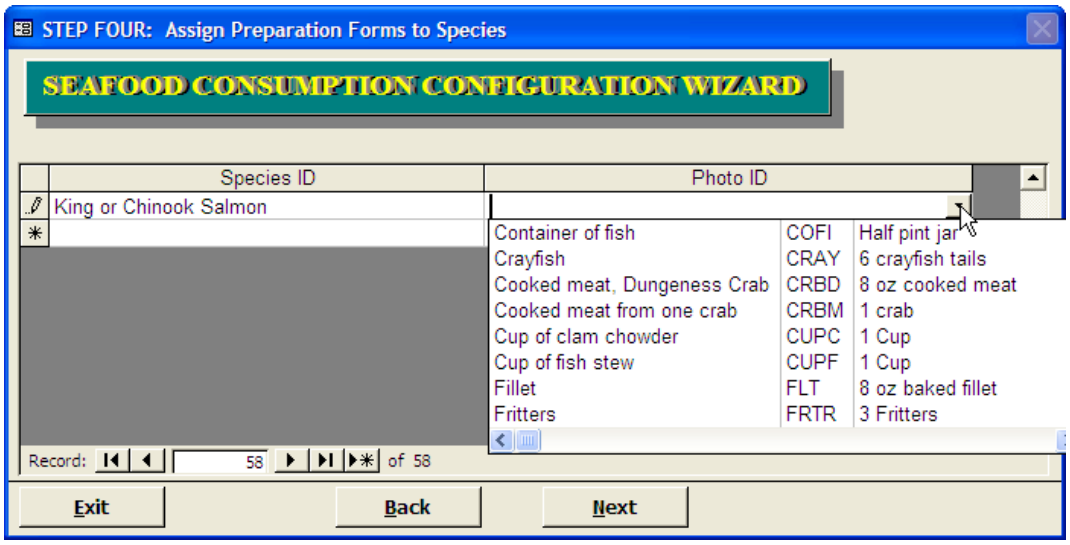


Figure 33. Select photograph in second column

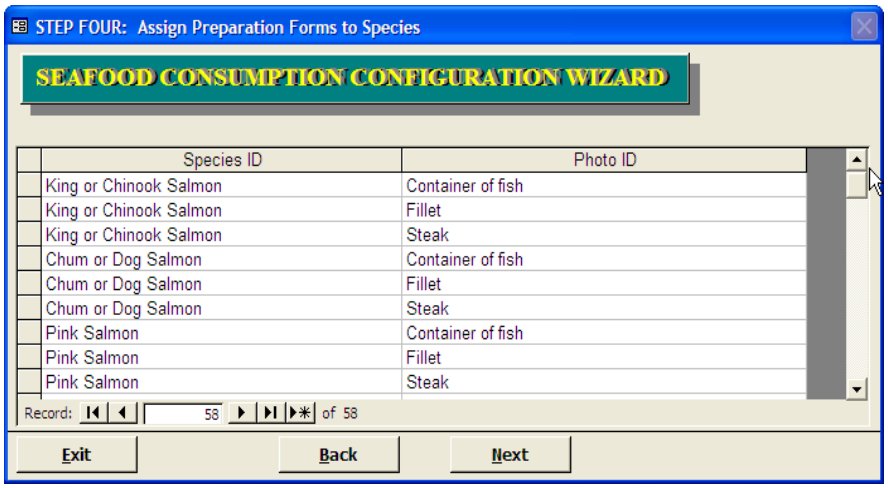


Figure 34. Photographs assigned to species in Quinault CAPI

In the above example, three photographs are assigned to Chum or Dog Salmon. The Respondent will be asked three questions if they respond that they eat Chum or Dog Salmon: if he or she eats Chum or Dog Salmon from 1) a container, 2) as a fillet, and 3) as a steak. The three photographs represent a standard portion size of these three prepared forms. During the CAPI, the interviewer will be able to record answers that reflect the Respondent's inability or unwillingness to select a portion model.

Press the **Next** button to proceed to the next configuration screen.

On the fifth datasheet screen (**Figure 35**) you define unique shellfish parts that will be asked for at least one shellfish food item. The system will assign a unique *Part ID* for each *Part Description* defined as shown below. Note that the part "Other Part" is pre-defined in the datasheet and will appear in the first row when you first open this datasheet.

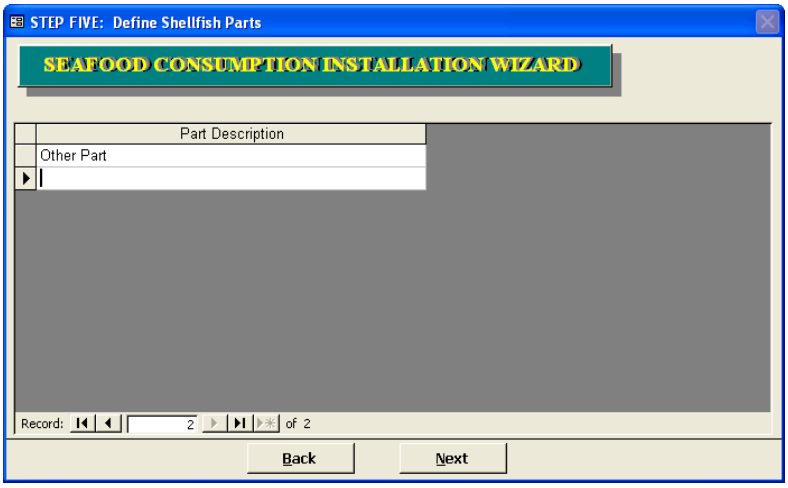


Figure 35. "Other Part" is defined by the application

Figure 36. Completed parts datasheet for Quinault CAPI

Press the **Next** button to proceed to the next configuration screen.

The sixth datasheet screen (**Figure 37**) is where you specify which shellfish parts are asked about for each species of shellfish. Shellfish parts are asked about in the consumption rate sections of the survey for adults and children. Choose a *SpeciesID* from the drop down box in the left field and an associated *Part ID* in the right field's drop down box.

Figure 37. Select shellfish Species ID in first column

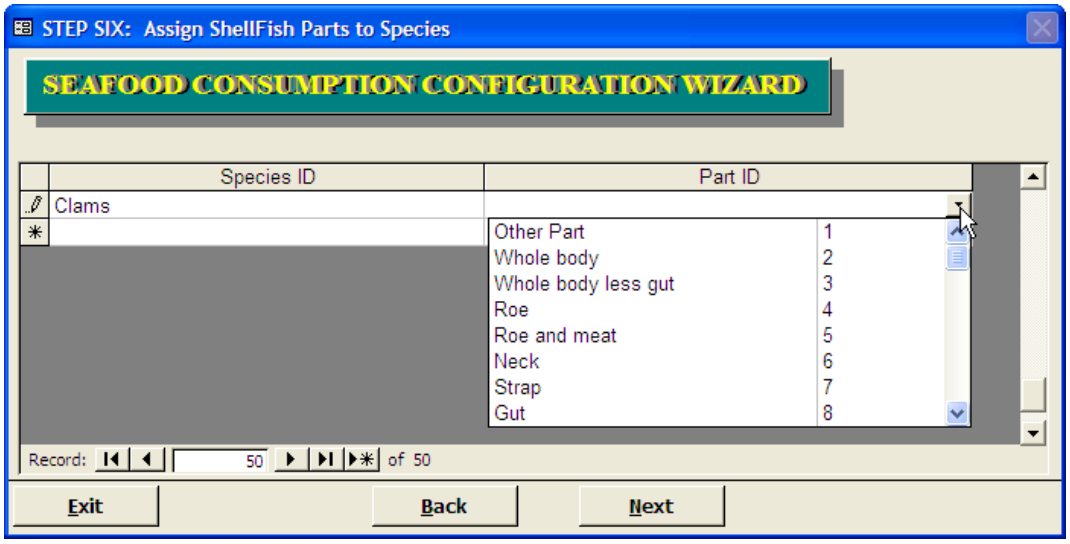


Figure 38. Select parts in second column

In the below example, five parts are asked about for Razor Clams in the CAPI. The Respondent will also have an opportunity to provide percentage eaten information for another part not specified by choosing "Other Part", though the Supervisor does not have to make that association on this datasheet.

Figure 39. Completed parts to ask for Quinault CAPI

Press the **Next** button to proceed to the next configuration screen.

On the seventh and last datasheet you will define the Shellfish harvest and Finfish catch locations in your geographic area. These values will be used to populate the drop down box for location questions in the CAPI. In the CAPI, if the Respondent reports a location not included on the list, the Interviewer will be able to choose location "OTHER" and be prompted to specify the name of the location provided by the Respondent.

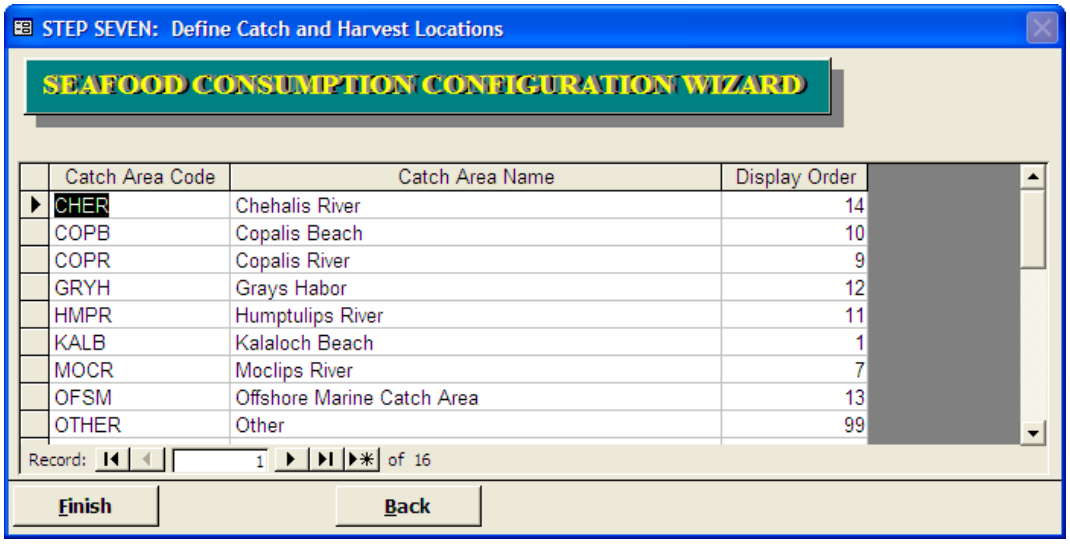


Figure 40. Locations of interest for Quinault Nation

Press the **Finish** button at the bottom of the screen to exit the configuration screens. You may modify the configuration datasheets as many times as you would like provided you have not published a final version of the CAPI.

CAPI Draft Version

Once the Supervisor has completed the seven configuration datasheets and would like to view the configured CAPI, he/she will choose the ***Publish Draft*** from the **Configure** menu item.

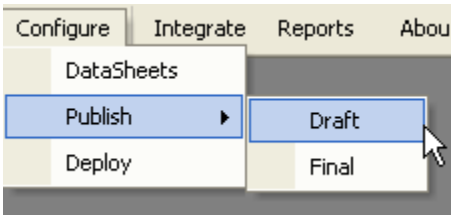


Figure 41. Publishing a draft version of the CAPI

Because this a new database again, you will have to "**Stop All Macros**" as instructed on pages 3-4.

Choosing this menu option updates a system table that tracks the versions of the CAPI. The draft version of the CAPI will then be started. Any previous draft version, including test interview data, will be overwritten. Exit from the CAPI software to return to the Supervisor Module. See the Interviewer's Guide for instructions on how to run the CAPI and exit.

You may publish and run the draft CAPI as many times as you like until you publish a final version. If after publishing a final version you try to publish a draft again you will see the error message shown in **Figure 42**.

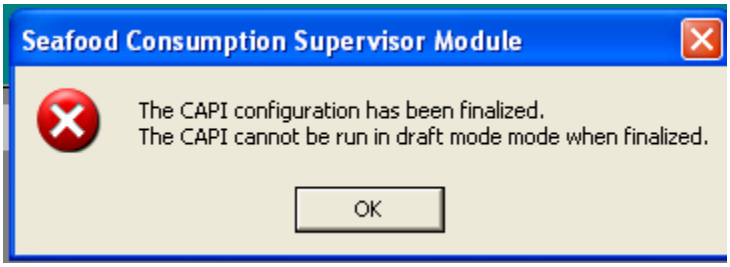


Figure 42. Error citing the Draft Publish options are no longer available

CAPI Final Version

Once the Supervisor is satisfied with the CAPI version, he/she will choose to publish a final version. It is up to the Supervisor to ensure the configuration is complete, accurate, and logical so that the CAPI performs effectively and as intended. Choose the ***Publish Final*** menu options.

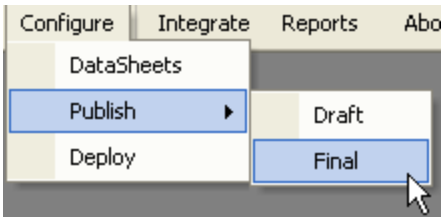


Figure 43. Publishing a final version of the CAPI

Choosing this menu option updates a system table that tracks the versions of the CAPI and makes note that the CAPI has been finalized and is ready to be deployed to sites. You will be asked if you want to review the CAPI software.

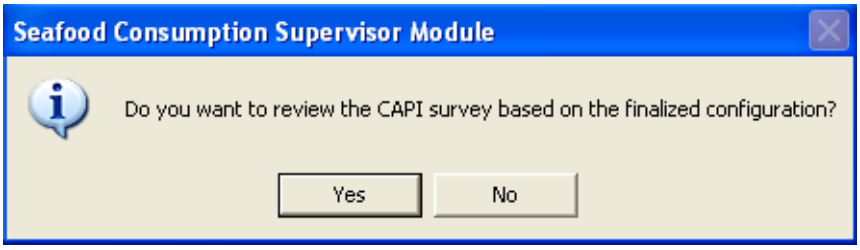


Figure 44. After final publication, Supervisor is asked to review the CAPI software

Press the **Yes** button to initiate the CAPI and press the **No** button to return to the menu.

This menu option creates a Supervisor's version of the final CAPI. The last draft version is now the same as the final draft. Should you want to run the final version subsequently, choose the Choose the ***Publish Final*** menu options. You will see the window in **Figure 44**. Test data can be saved in the final draft version. This Supervisor's version of the CAPI database will not be the version used to compile data from the sites upon import. The Supervisor, therefore, may run the final version and add test data without affecting the database that is deployed to the sites.

Once you publish a final version, you may no longer add new records to the Species Question Library through the main menu and modify text to existing records in the Species Question Library, and you may no longer access the configuration datasheets.

Deployment

Once a final version of the CAPI has been published, the Supervisor can deploy the software to the sites. Choose the menu option ***Deploy*** to initiate this function. You should have ready a USB storage device, also known as a thumb-drive or flash-drive.

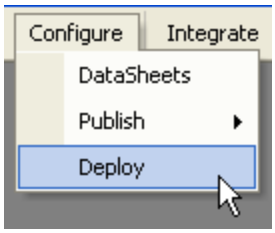


Figure 45. Deployment menu item

A window will be displayed asking you to provide the name of the site to receive the software.

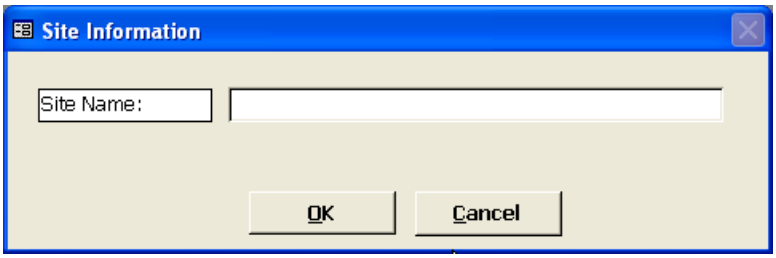


Figure 46. Prompt for site information

Pressing the **Cancel** button will return you to the main menu.

If the **OK** button is pressed, the module will create one zip file containing the software and database needed by the sites to install the CAPI locally. In addition, the Install.exe file to guide the interviewer through the installation process at the site will be copied to the USB device. The device can then be given to an interviewer to install the CAPI software on their PC. After entering the Site Name, you will be prompted to enter the location of where you would like the deployment files will be written.

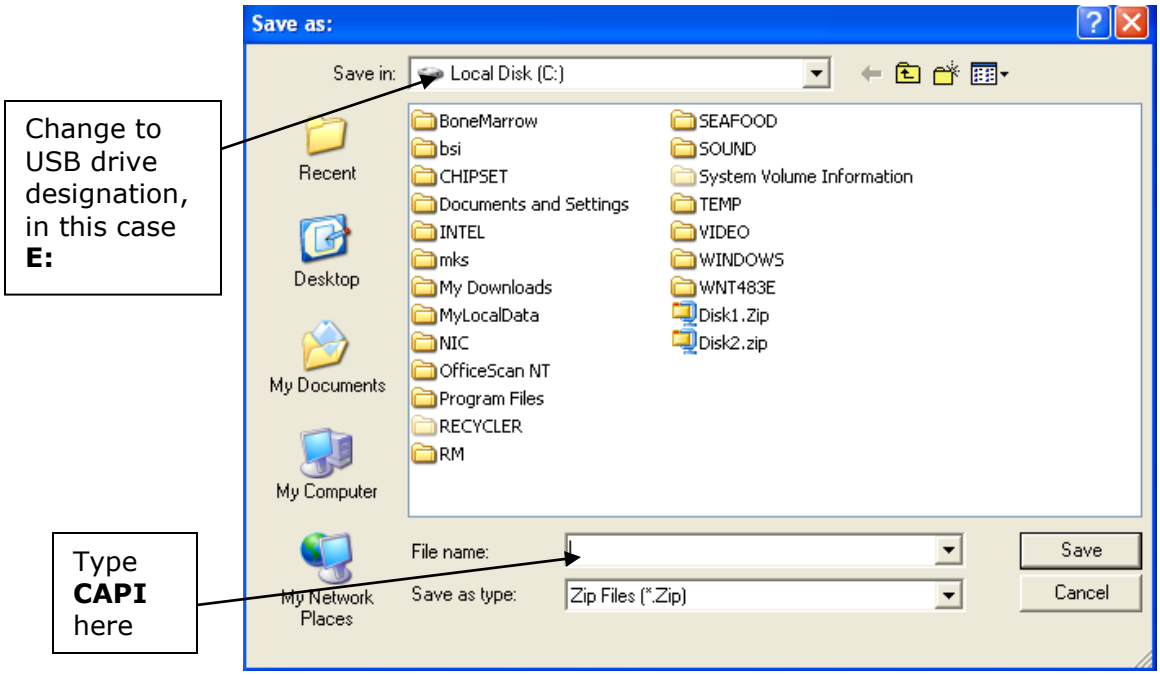


Figure 47. Prompt for deployment file location

Insert the USB device into a port (in **Figure 48** it is the "E" drive) to write the files to the device in the *Save In* field. Type in the file name **CAPI** and press the **Save** button.

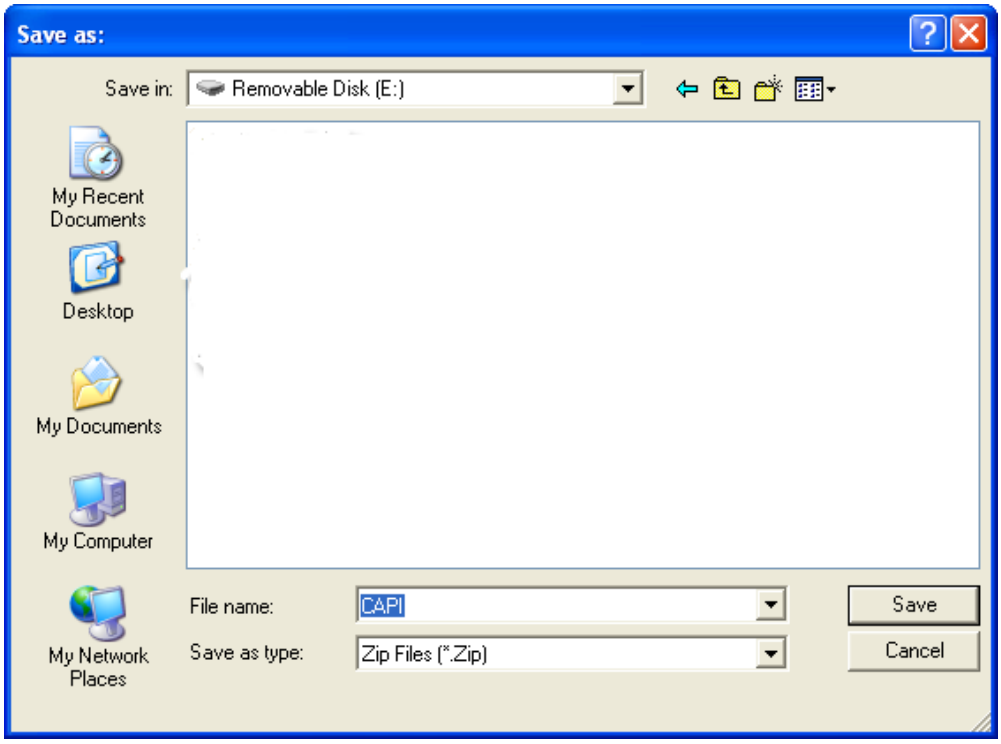


Figure 48. Prompt for Deployment file location

You will be presented with a warning that any installation files previously installed on the device in any folder will be deleted before the CAPI installation files will be written.

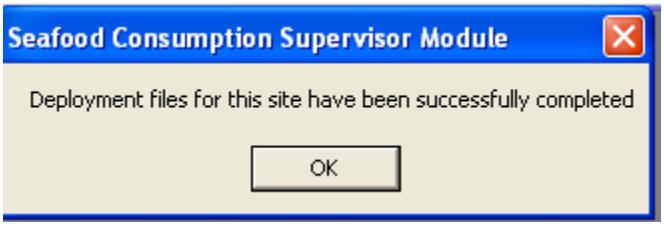


Figure 49. Prompt to determine success

Accumulation of CAPI Data from Sites

The Supervisor Module has the functionality to integrate interview data collected at the deployed sites. Accumulation of a site's data into the Supervisor's master version (which differs from the Supervisor's final version) is a three step process. First, the site specific *SeafoodSurvey_dat.mdb* file has to be imported, then the "Other Specifieds" recorded by Interviewers need to be translated to a *SpeciesID* that is part of the Species Question Library, and finally the data is added into the master database.

Import

Once an interviewing site has completed its interviewing, it will export its data to a USB device, possibly more depending upon the number of completed interviews. This process is explained in the Interviewer's User Guide. This device should be sent to the Supervisor with care.

Insert the USB device received from the site into the Supervisor's machine USB drive. The Supervisor will then import the data on the diskette(s) to his/her PC by choosing the ***Import*** menu item.

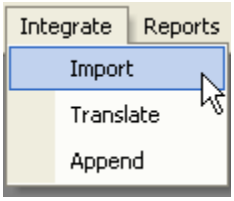


Figure 50. Import menu item

A window listing deployed sites is opened.

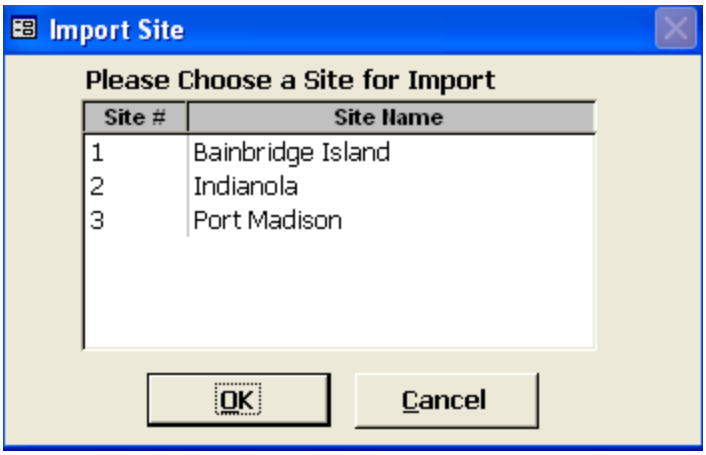


Figure 51. Choose a site to import

Highlight the site you would like to import. Press the **OK** button to start the import or press **Cancel** to return to the menu. If you continue to import, you will be asked to specify the location of the data to be imported.

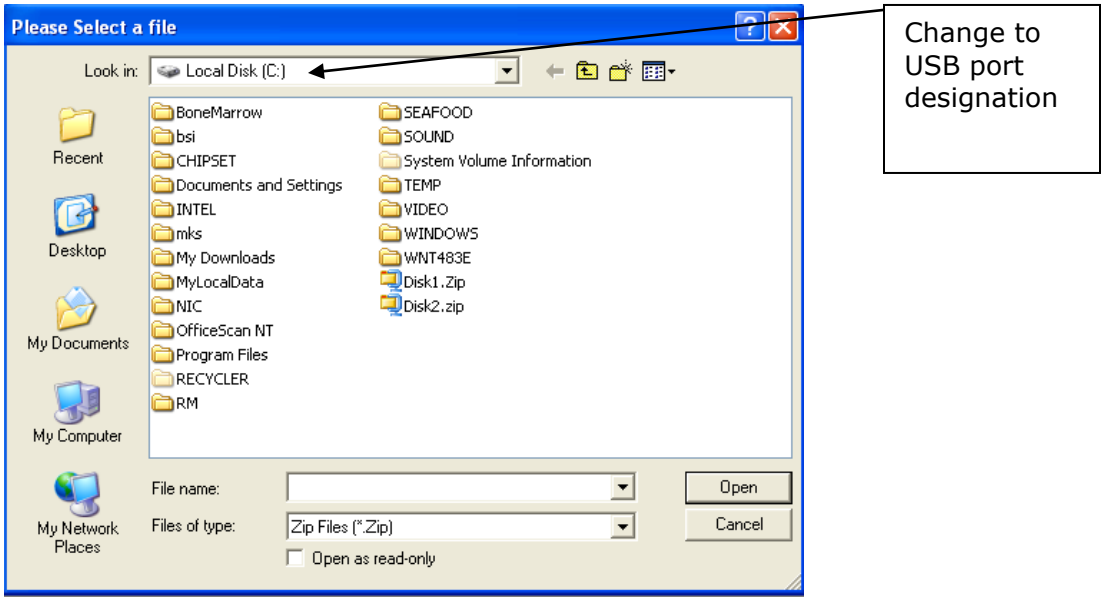


Figure 52. Specify the drive letter where the site's USB device has been inserted

Press the **Open** button. You should see a file with the name of **trans*siteno*.zip**, where ***siteno*** is the site number assigned by the system during deployment.

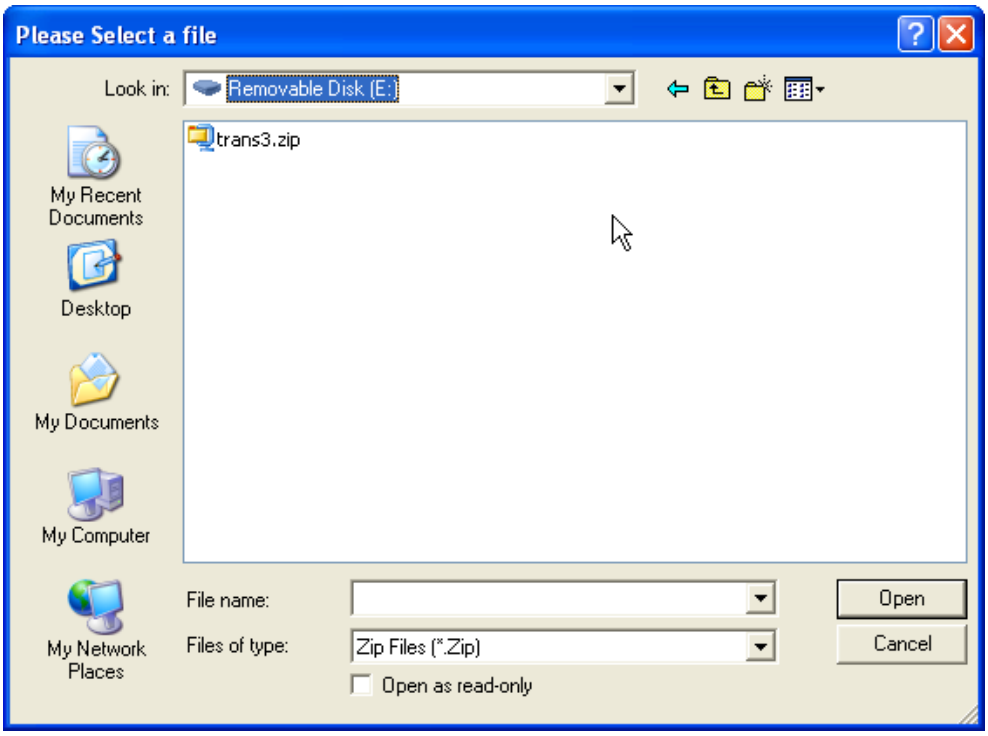


Figure 53. Highlight the file name with your mouse

The site number should match the number that corresponds to the site name chosen in **Figure 51**. Highlight the file with your mouse. Press the **Open** button. If the site number of the file name on the diskette does not match the site number of the site chosen in **Figure 51**, the error message in **Figure 54** will be displayed.

Figure 54. Error when file name site number does not match selected site number

Press **OK** to return to site selection screen in **Figure 51**.

If the site numbers match, the system will copy the file in compressed format to *C:\SupervisorSeafood\Import* and write an uncompressed version to *C:\SupervisorSeafood\Translate*. The uncompressed version will have a file name of **trans*siteno*.mdb**, where ***siteno*** is the site number assigned by the system during deployment. You will be notified when the process has completed.

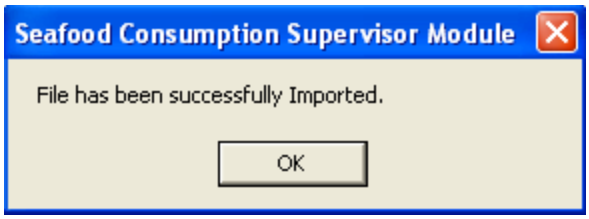


Figure 55. Import has been successful

The selected site will no longer be available for import and will be absent from any future site selections list displayed in **Figure 51**.

The site's CAPI data is then ready to be prepared for accumulation.

Translate

During the interview at the end of the 24-hour Recall and following annual Shellfish consumption, the Respondent is asked if he/she ate or eats any other seafood that was not asked about during the CAPI. These are referred to as "Other Specifieds" and have been categorized to Group "X" in the site database. If a site's database contains "Other Specifieds", the site database cannot be added to the master database.

The ***Translate*** menu item will allow you to handle the other seafood items reported by a Respondent.

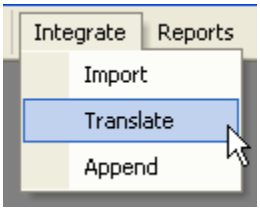


Figure 56. Translate menu item

Other reported seafood items specified are assigned a temporary *SpeciesID* during the administration of a CAPI. They must be translated to a *SpeciesID* that exists in the Species Question Library. Having the same *SpeciesID*s across sites facilitates analysis. Note, translated items remain in Group ID "X" and do not appear on any of the pre-defined reports.

Choose ***Translate*** to display the names of the sites that have had their data imported.

Figure 57. Listing of sites ready for translation

Use the mouse to highlight a site name and press the **OK** button. The **Cancel** button will return you to the menu. If **OK** is pressed, a list of other specified seafood will be displayed.

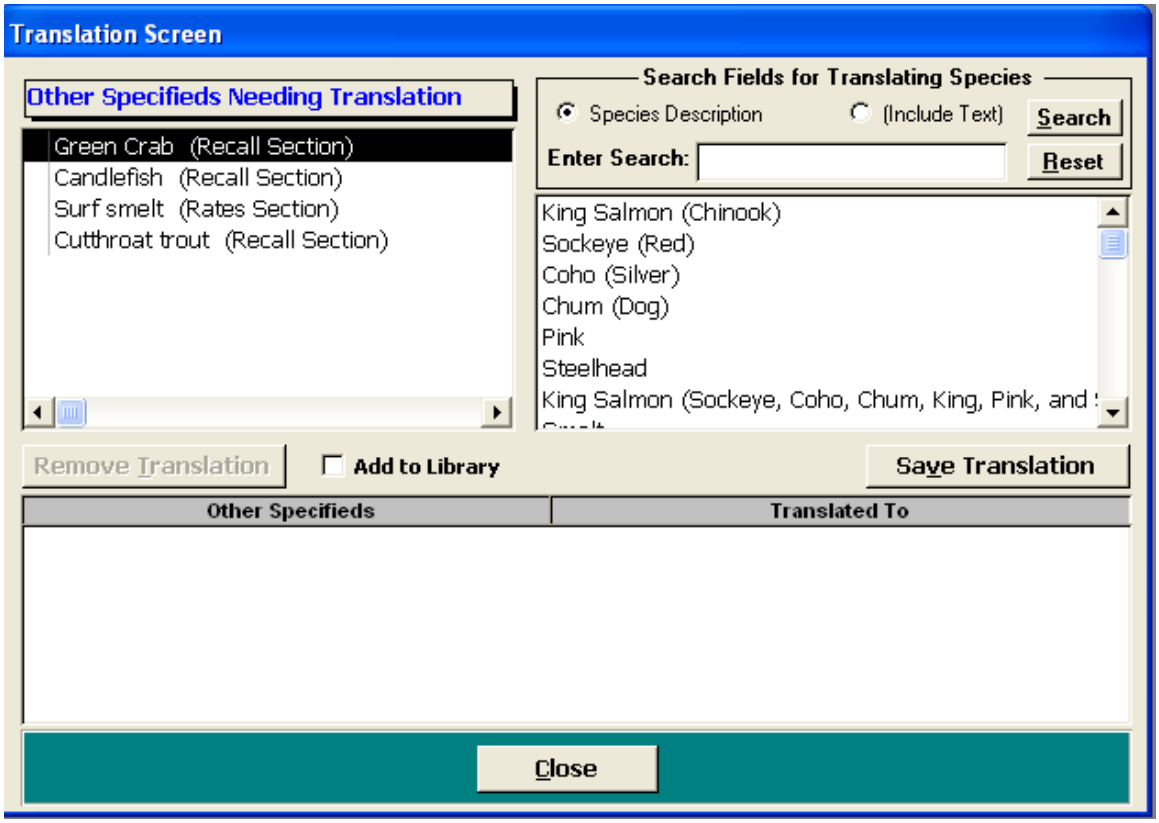


Figure 58. Other seafood items to be translated

The species list on the right hand side provides a list of species from the Species Question Library. Search functionality has been provided to more easily find a

species from the master library for the "translated to" value. To translate "Surf Smelt" use your mouse to highlight the item in the upper left pane. To see what "smelt" species currently exist in the Species Question Library you can search the library by entering in "smelt" in the field after the "Enter Search" label.

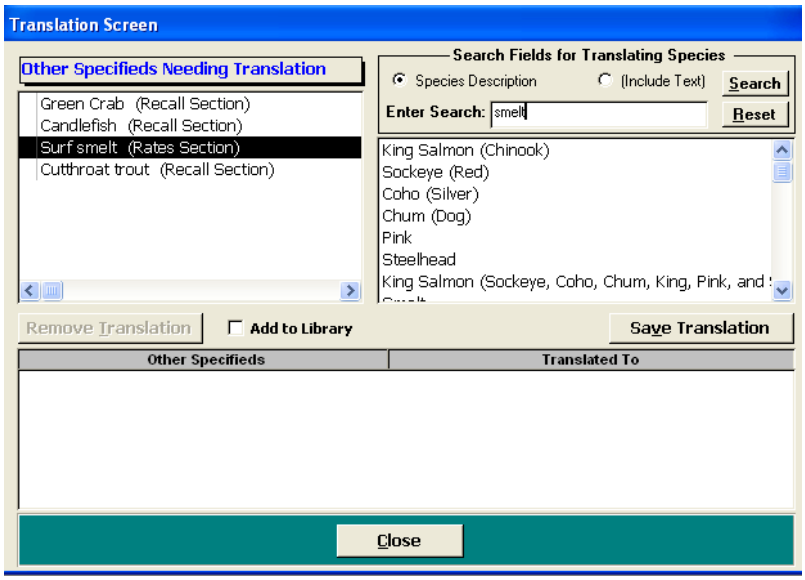


Figure 59. Search for "smelt" in Species Question Library.

Press the **Search** button. The results are below.

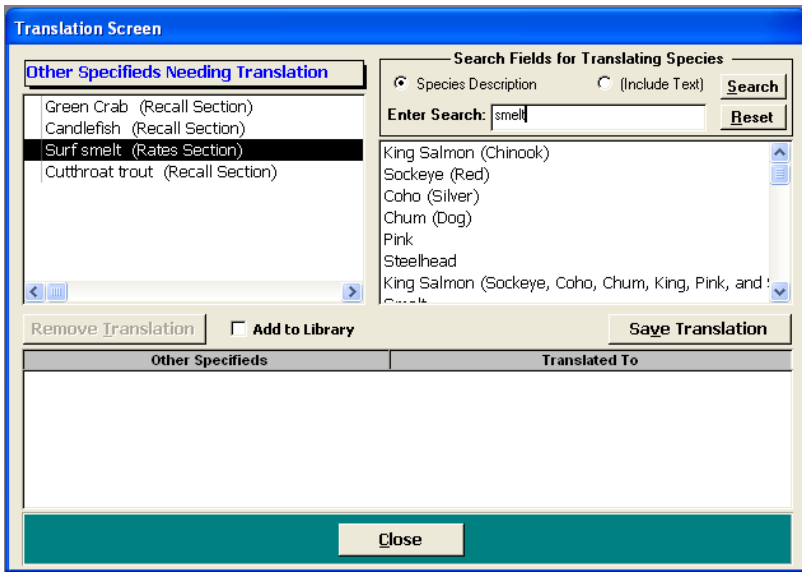


Figure 60. Search results for "smelt" in Species Question Library.

If you choose to translate "Surf smelt" to "Smelt", highlight "Smelt" in the upper right pane and press the **Save Translation** button. The translation will appear in the lower half of the window and an "x" will appear to the right of "Surf smelt".

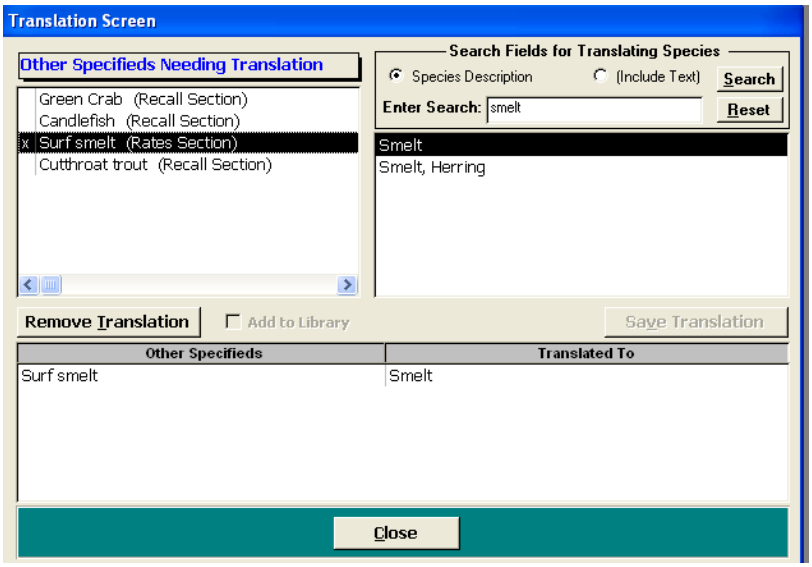


Figure 61. "Surf smelt" has been translated to "Smelt"

Highlight any translated "Other Specified" in the upper left to display the translation in the lower half of the window. If you would like to delete the translation, highlight the record in the bottom half of window and press the **Remove Translation** button. The record at the bottom will disappear as will the "x" to the left of the item in upper left box.

You must translate each item. You cannot translate two "Other Specifieds" to the same *SpeciesID* so as not to comprise database integrity. If you cannot find an appropriate translation in the existing Question Library, you can add a new entry to the library. Highlight the "other specified" to be translated, click the "Add to Library" checkbox (**Figure 62**). Press the **Save Translation** button. The text will be added to the Species Question Library under *SpeciesDescription* and a unique *SpeciesID* will be assigned by the module. The item will be translated to itself and appear at the bottom of the screen under items translated.

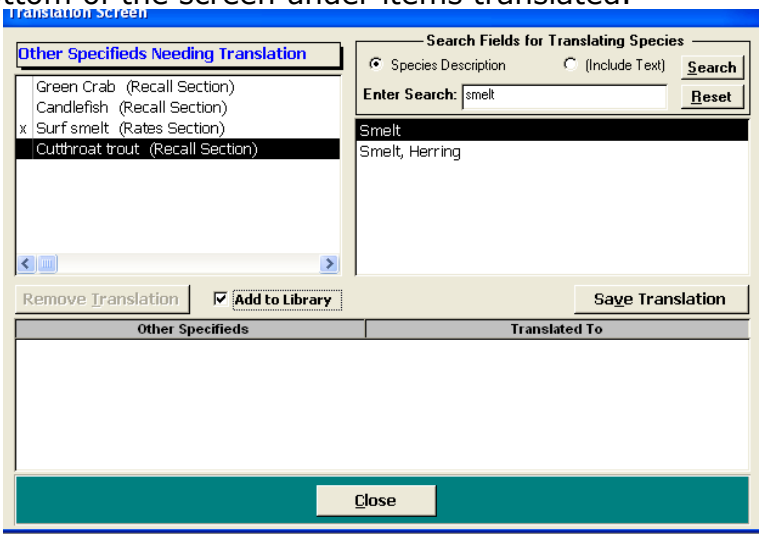


Figure 62. Add Cutthroat trout to the Species Question Library

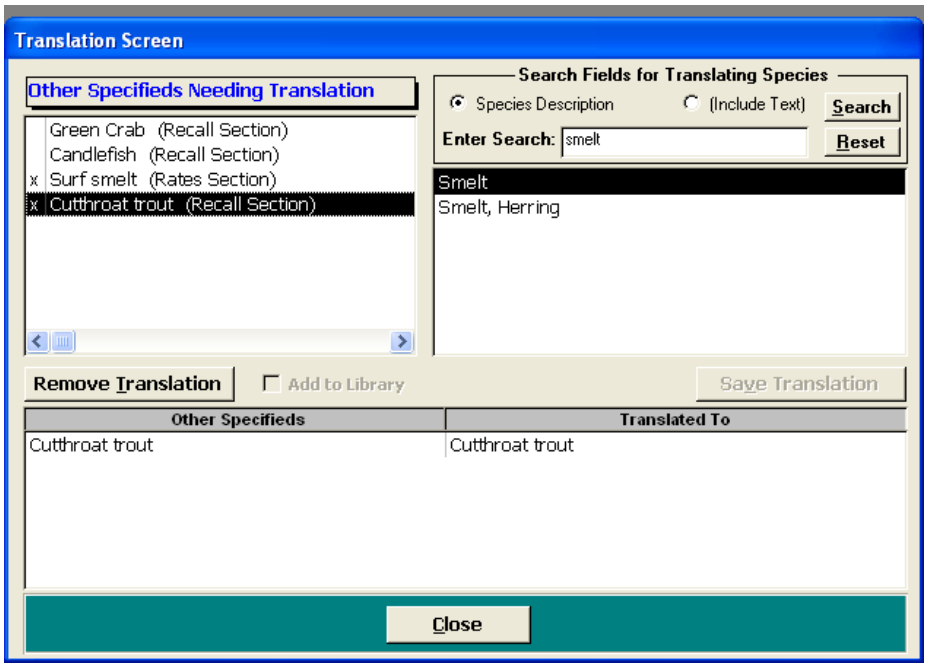


Figure 63. Cutthroat trout has been added to the library

You can choose the ***Translate*** menu option multiple times as long the site's data has not been appended that site's data into the master database. You will not be able to append site data if there are any other species that have **not** been translated.

Append

Once there are no items to be translated, the site data can be added to the master database.

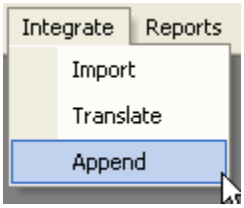


Figure 64. *Append* menu option

When you choose ***Append***, the module will display the names of the sites that have had their data imported.

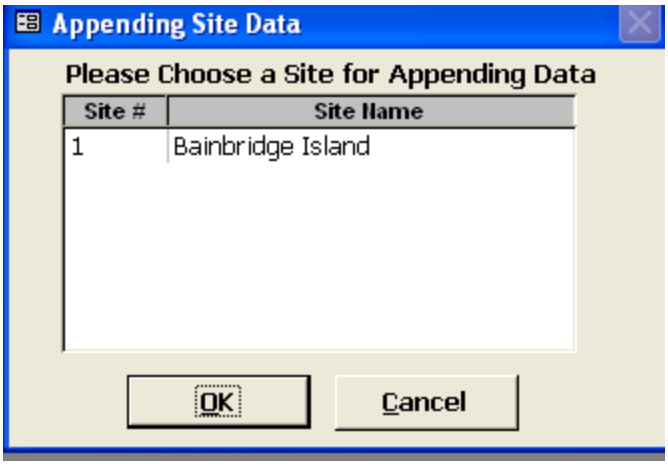


Figure 65. List of sites ready to be appended to master database

Highlight a site name and press the **OK** button. The module will then append all site data from the selected site to the master database and display a successfully completed message as shown in **Figure 66**.

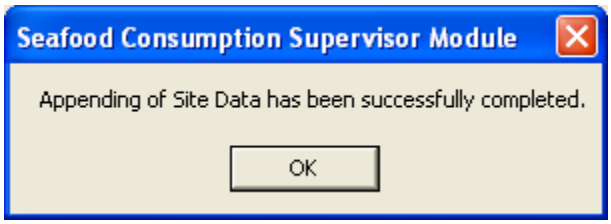


Figure 66. Notification of successful integration

Reports

In the Supervisor Module, there are six statistical reports defined in the application software. These reports can be found under the ***Reports*** main menu item. Descriptive statistics can be generated for 24-hour recall and seasonal/annual rate data at the species item level, at the group level, and overall. Overall and group statistics by gender for 24-hour recall and rate data are also available.

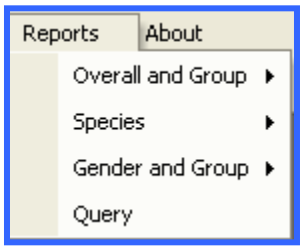


Figure 67. *Reports* menu options

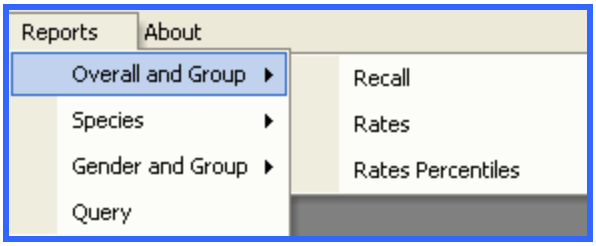


Figure 68. *Reports* sub-menu options

Query

Under the ***Reports*** menu item, there is a ***Query*** option that offers a structured method of querying the adult consumption rates ascertained in the Rates section of the CAPI on pre-defined selection criteria. These rates are stored in *TblRates* table. You will be taken through a series of seven screens that control what rates data will be presented in the results. You can choose to export the result to a comma delimited file or to execute a report using the results as the basis. The report includes the count of the number of subjects included in the subset (summed across seasonal records), mean of the annual gram amount consumed, the standard deviation and variance of the amounts as well as percentile values.

Choosing this menu item will bring up the first window in a series that asks you to select whether you want to look at species asked in the Finfish rates section or whether you want to look species asked in the Shellfish rates section of the CAPI. Once, you have made your selection, the relevant groups are populated in the bottom left window.

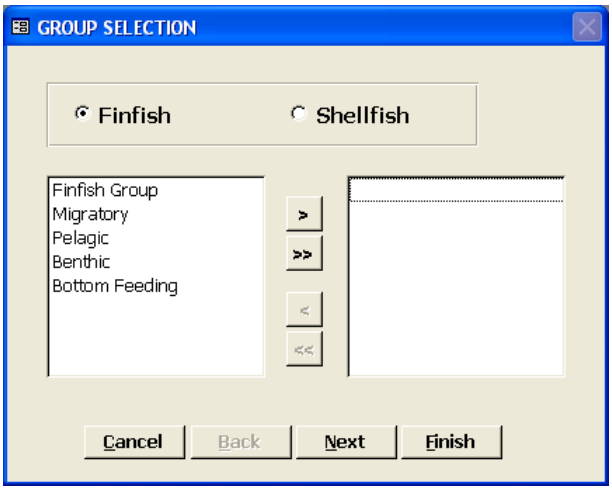


Figure 69. First query filter screen

Use your mouse to highlight a group or groups of interest and move over to the right box using the operators.

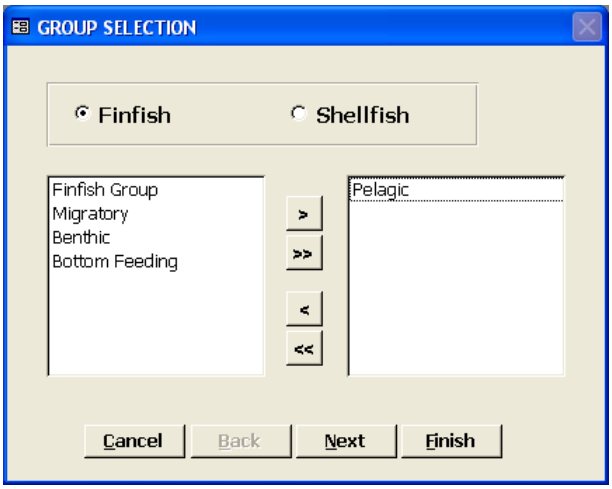


Figure 70. Selection of group of interest

You can press Finish at any time to move to the output option screen where you can either export the queried data or preview the statistical results in a report format.

Press the **Next** button to move to the next query screen or choose the **Cancel** button to return to the main menu.

This next screen will ask you if you would like to see only species questions related to migratory fish, resident fish, or both. The default is both.

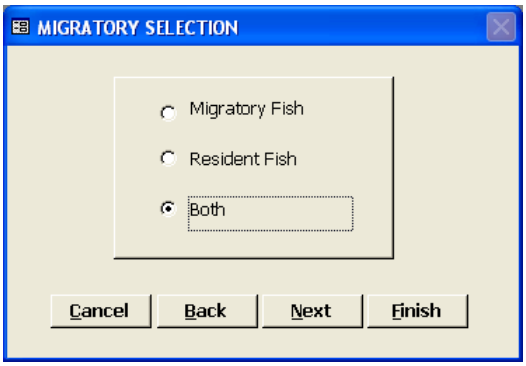


Figure 71. Second query filter screen

Press the **Next** button to move to the next query screen or choose the **Cancel** button to return to the main menu.

This next screen will display all of the locations defined during the configuration, for the group(s) that you have chosen on the first screen. Select locations of interest from the left pane and move to the right pane.

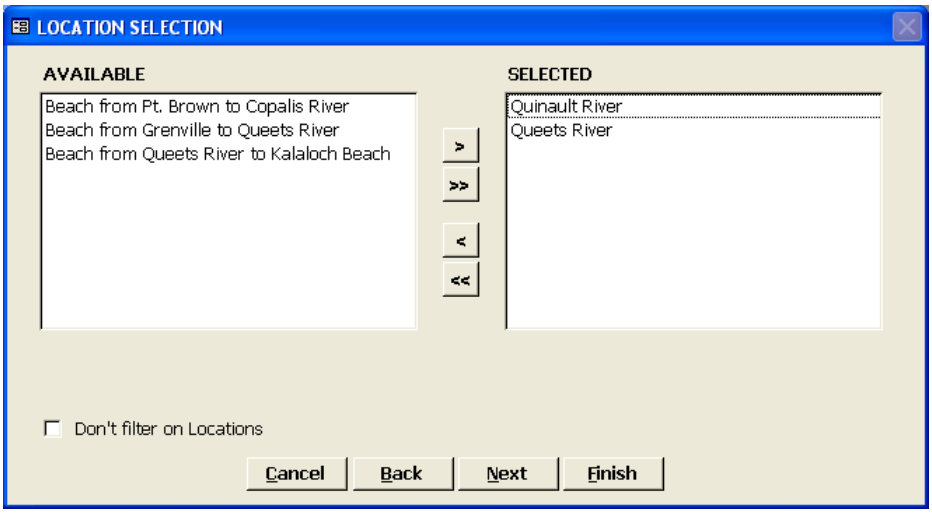


Figure 72. Selecting location on third query filter screen

Press the **Next** button to move to the next query screen or choose the **Cancel** button to return to the main menu. If you do not wish to choose a location but would like to keep going through the query steps, choose the "Don't filter on locations" checkbox and press the **Next** button.

This next screen will display all of the defined sources of seafood, for groups satisfying the criteria specified in the first two screens. Select the source or sources of interest from the left pane and move to the right pane.

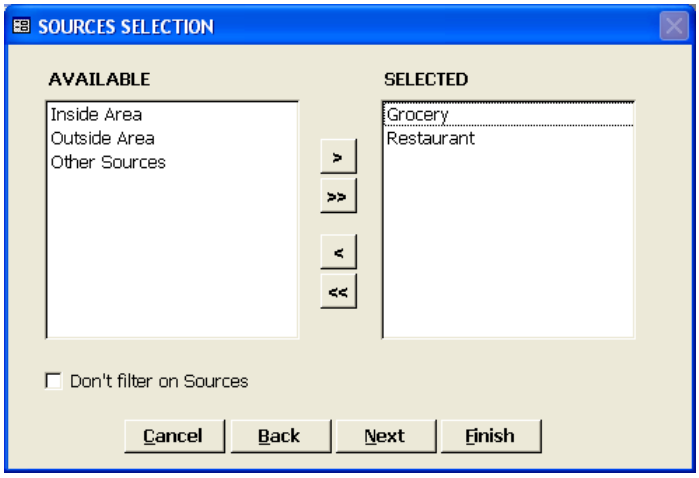


Figure 73. Selection of sources on fourth query filter screen

Press the **Next** button to move to the next query screen or choose the **Cancel** button to return to the main menu. If you do not wish to choose a source but would like to keep going through the query steps, choose the "Don't filter on sources" checkbox and press the **Next** button.

This next screen will display the species questions that match the criteria specified in the previous four query screens. Use your mouse to highlight a species question or species questions of interest and move over to the right box using the operators.

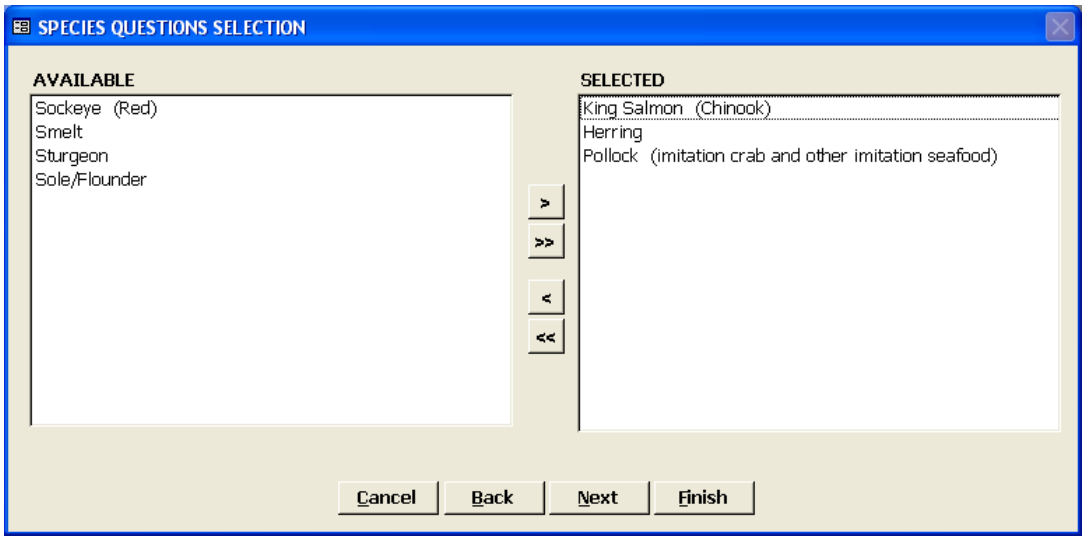


Figure 74. Selection of species of interest on fifth query filter screen

Press the **Next** button to move to the sixth query screen or choose the **Cancel** button to return to the main menu.

This screen asks whether you want to include in your subset subjects that consumed each of the species selected in the previous screen or subjects that consumed at least one of the selected species. Subjects that consumed all species are called All Consumers while subjects that consumed at least one species are called Ever Consumers.

Figure 75. Selecting consumer type on the last query filter screen

Press the **Next** button to move to the last query screen or choose the **Cancel** button to return to the main menu. Select what you would like to do with the results.

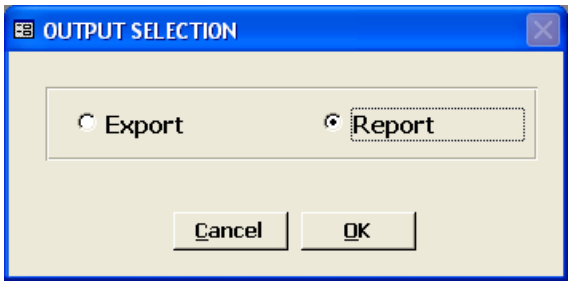


Figure 76. Selecting output format for query results

The Export option will save the results to a comma delimited file that can then be imported into software such as Excel or SAS. There will be one record for each seafood item reported consumed and stored in *TblRates* that meets the subsetting criteria provided. For Finfish species seasonal rates are asked for in the CAPI. If the Respondent was able to provide both in and out of season rates, there will be two records, one for each time period, in the file for that species. In addition to the fields in *TblRates*, the export file will contain a calculated annual consumption amount based on reported frequency and raw portion size. You will be prompted for a location and file name.

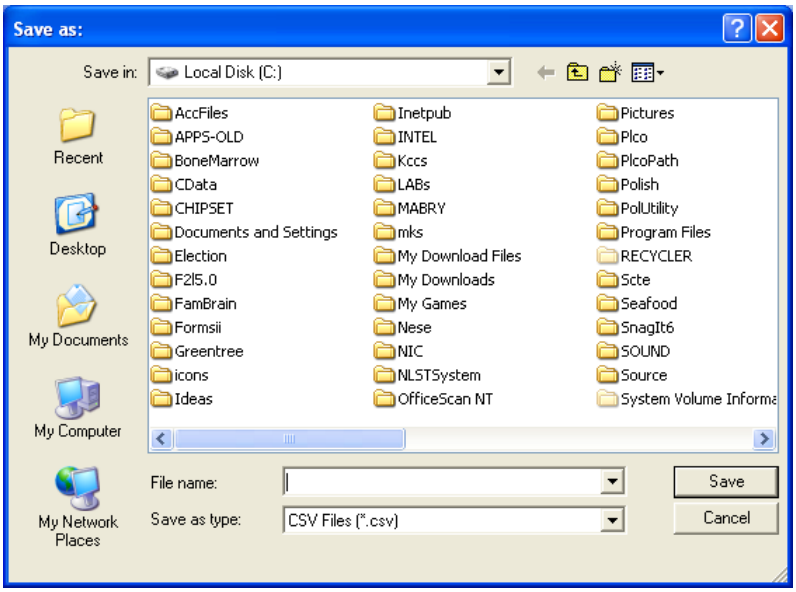


Figure 77. Providing a location and file name for export file

Once you provide a location and file name and press the **Save** button, the file will be created.

Pressing the **Cancel** button will return you to main menu.

If you choose the Report option in **Figure 76**, a report will be generated and displayed in preview mode. The report will contain the count of the number of subjects included in the subset (summed across seasonal records), mean of the annual gram amount consumed, the standard deviation and variance of the amounts as well as percentile values. You can choose to print the report or close the report.

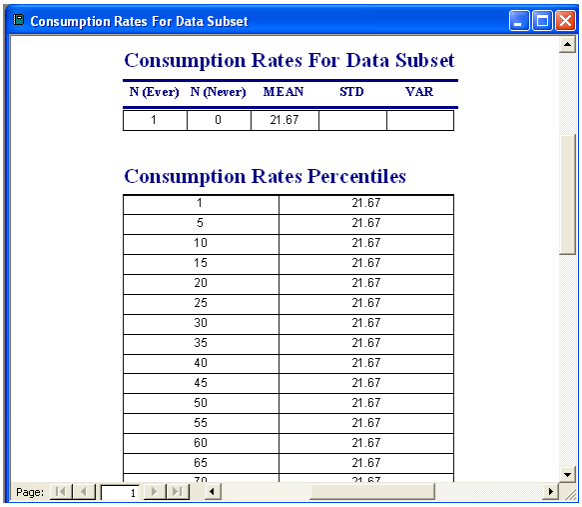


Figure 78. Report results from query search

The Fish Consumption Survey Tool (FCST)
CAPI Survey
Description of the Interviewer Software

The Fish Consumption Survey Tool (FCST) is administered by computer assisted personal interview (CAPI) software developed in MS Access. The Suquamish Nation Seafood Consumption Survey, a previously used, hardcopy, interview administered questionnaire, is the basis of the design of the CAPI instrument. The screen shots included below reflect the survey as configured for the Quinault Nation in March 2006.

The software consists of two Access 2010 databases: one database (*SeafoodSurvey_dat.accdb*) contains the data collected during interviews and lookup tables (i.e. tables that contain the various possible answers to interview questions) while the second database (*SeafoodSurvey_app.accdb*) contains the application code and objects such as the queries, forms, and reports. These databases will be copied to the local folder *C:\Seafood* during installation.

Computer Requirements

The designated Interviewer computer must have a Pentium III or higher processor. It must have at least 128 MB of memory though 256 MB is preferable. The hard disk should have at least 60 MB available space. The computer's operating system should be Windows 7. Access 2010 or higher should be installed on the computer prior to installing the CAPI software.

Setup

The CAPI software has been customized for your study so that seafood typically consumed by the population are included in the survey. This customization has been completed on the Supervisor PC and passed on to sites on a USB storage device (also known as a flash-drive or thumb-drive). Interviewers are responsible for installing the CAPI on their designated PCs referred to as Sites.

To install the CAPI, insert the USB device into your laptop's USB port. For these instructions it is assumed the port is in the "**E:**" drive. Open Windows Explorer. Windows Explorer may be found in many places on your PC and it varies from PC to PC. In Windows 7, typically, you can find it by pressing Start, choosing Programs, choosing Accessories, and choosing Windows Explorer. Use Windows Explorer to expand the "**E:"** drive (or the USB port/drive used on your PC). You will see:

.

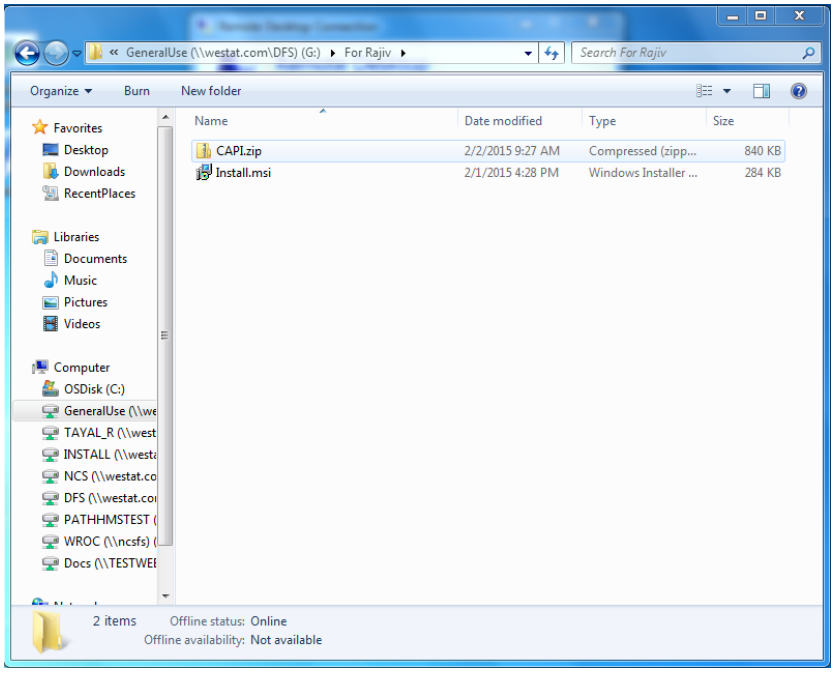


Figure 1. Opening Windows Explorer

Double-click on the file ***Install.msi***. The installation procedure will copy the necessary software files to the local folder *C:\Program Files(x86)\Seafood Survey* unless otherwise specified during installation.

To begin the second part of the installation, double click on the file ***InstallApp.accdb*** in the C:\Program Files(x86)\Seafood Survey folder as shown in **Figure 2**.

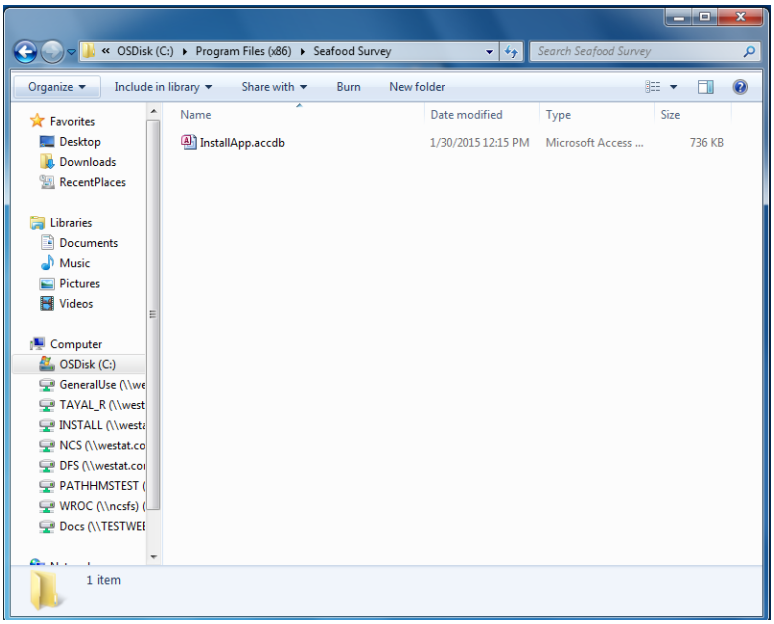


Figure 2. Double click on *InstallApp.accdb*

Click on the Cancel button then Click on the Enable Contents button. Select the **Run Install** button to continue installation.

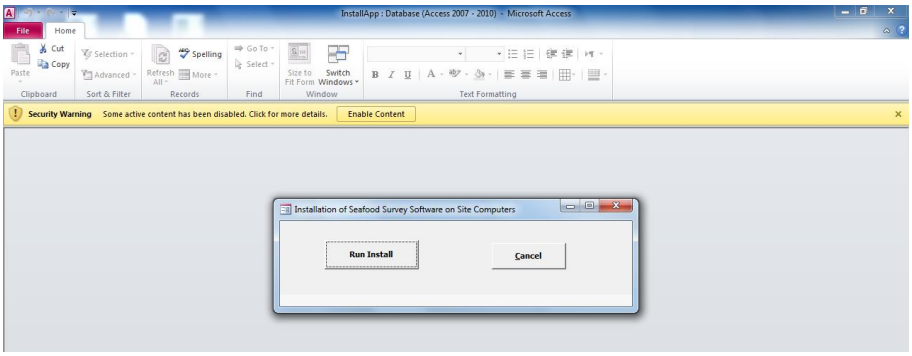


Figure 3. Select Cancel

Press **OK** to continue at the next prompt.

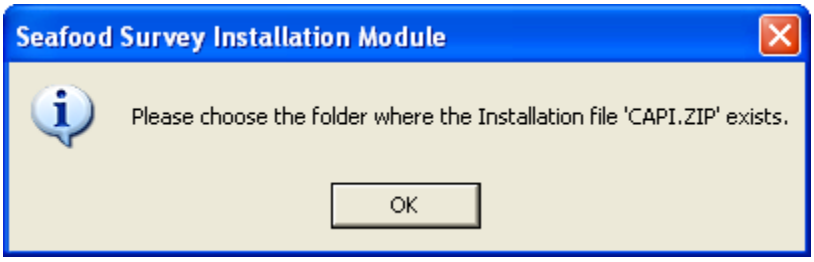


Figure 4. Press OK

Once the screen below is presented, change the drive designation to the USB port (drive "**E:**" in Figure 5).

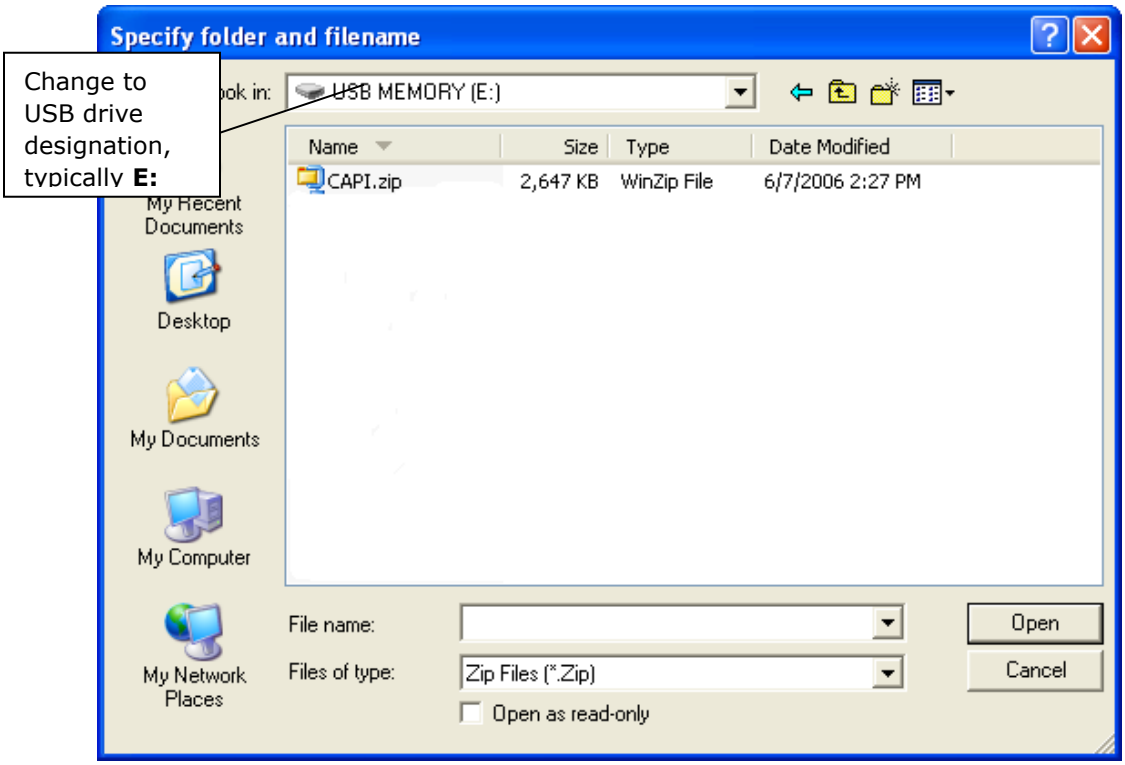


Figure 5. Specify the USB drive

Highlight the file ***CAPI.zip*** and select the **Open** button.

You will be notified of the successful installation.

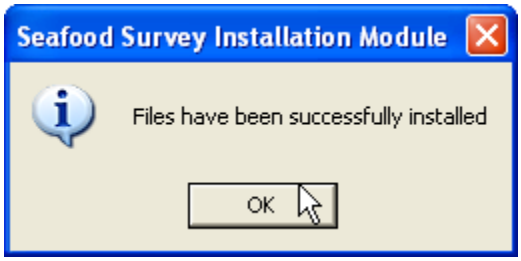


Figure 6. Notification of successful installation

You will need to copy a shortcut file to your desktop to allow for easy access to the CAPI software. While there are a few different ways of doing this, this guide will describe one way. Using Windows Explorer, go to the *C:\Seafood* folder. Right click on **Shortcut to Seafood Survey** file and choose "Copy" from the menu.

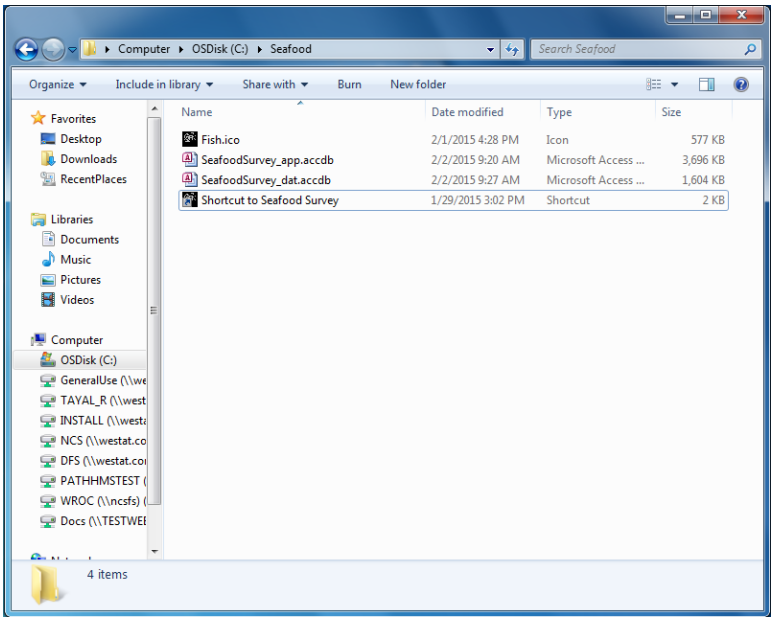


Figure 7. Shortcut file in the folder

Minimize all opened windows by choosing the minimize window control at the top right corner of a window  . You should then be on the desktop. Right click anywhere on the desktop to display a function menu and choose "Paste".

Figure 8. Shortcut file pasted to the desktop

If you prefer, you can use your mouse to drag the icon to anywhere on the desktop.

When start the survey for the first time a prompt to "Enable content" will appear (Figure 9)

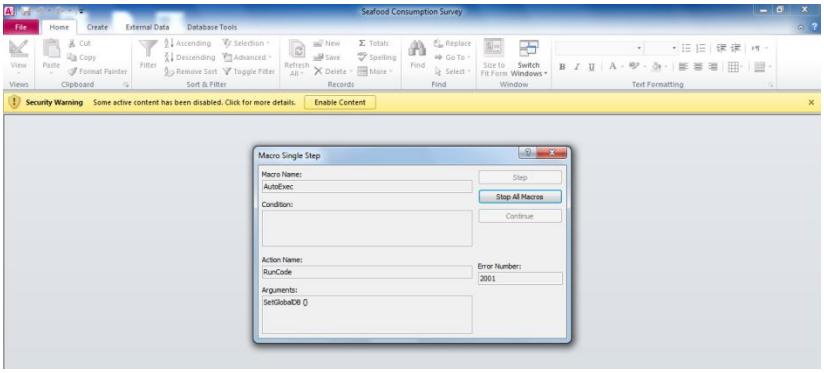


Figure 9. Enable Content prompt

Select the **Stop All Macros** button. Then select the **Enable Content** button in the yellow header.

You will be presented with the following screen:

Figure 10. Interviewer prompt

Enter the Interviewer name that has been assigned to you and select **Ok**.

Groups and Species

The CAPI structure and flow are in large part based on the definition of seafood groups and the species assigned to those groups of interest to the study. The Supervisor defines the groups when customizing the CAPI. Examples of groups are: 1) Migratory Finfish 2) Bottom-feeding Finfish and 3) Shellfish. Specific seafood types are assigned to these groups by the Supervisor. For example, Sole, Flounder, and Rockfish might be assigned to the group Bottom-feeding Finfish while Clams and Crabs would be assigned to Shellfish.

Interview Structure/Flow

The CAPI is organized into the following sections:

Startup In this section, the Interviewer will introduce the survey to the Respondent, record the date and start time of the interview, record the Respondent's gender, and determine if the Respondent has previously completed a survey within the designated time period. The survey is designed to allow for multiple administrations to a Respondent. A survey is uniquely identified by a Subject ID. A Subject ID is made up of the Study number, Interviewer name, Person ID, and Interview number. The Subject ID appears at the top of every page. Person ID uniquely identifies the Respondent and is assigned outside the survey. The Interviewer enters the Respondent ID. The system will calculate the survey interview number by incrementing the last survey number for that Person ID. If it is the first survey, the interview number is 1. The survey system will construct the full Subject ID. For the Subject ID 574-00-12345-1, the Study number is "574", the Interviewer name is "00", the Person ID is "12345" and the survey interview number is "1".

24-Hour Recall In this section, the Respondent is asked if certain seafood items were eaten yesterday, the day before the interview. The Respondent is also given the opportunity to report on other seafood items consumed yesterday but not already asked about.

Finfish In this section, the Respondent is asked about finfish consumption over the course of a year or in and out of season. Information included in this section of the survey includes: consumption rates and amounts, fish sources (e.g. grocery store, restaurant, and catch area), and parts of the fish eaten. This section is divided into multiple subsections, one for each group of finfish defined by the Supervisor. Starting with the first group, the Respondent is asked about consumption of species when the fish are in season and when the fish are out of season. If the Respondent is unable to report by season, the Respondent is asked about annual consumption patterns. Quantifying seafood consumption is done through the use of photographs depicting typical portion sizes of seafood prepared in different forms such as steak or in chowder. Once the Respondent provides consumption information for the first group of fish, the Respondent is asked to report on what percentage of the time he/she eats specific parts, considering all the fish from the first group typically eaten. The Respondent is also asked to report on from where the fish consumed was generally obtained. The locations have been predefined by the Supervisor though the Respondent has the option to report a location not defined by the Supervisor. If the fish was reported to have been caught in or outside the area of interest, the Respondent is further prompted to provide the location of where the fish was caught. A study may choose not to gather information about harvest or catch locations outside the area of interest. The Interviewers can skip the prompts for outside area information by answering "No" at the following prompt.

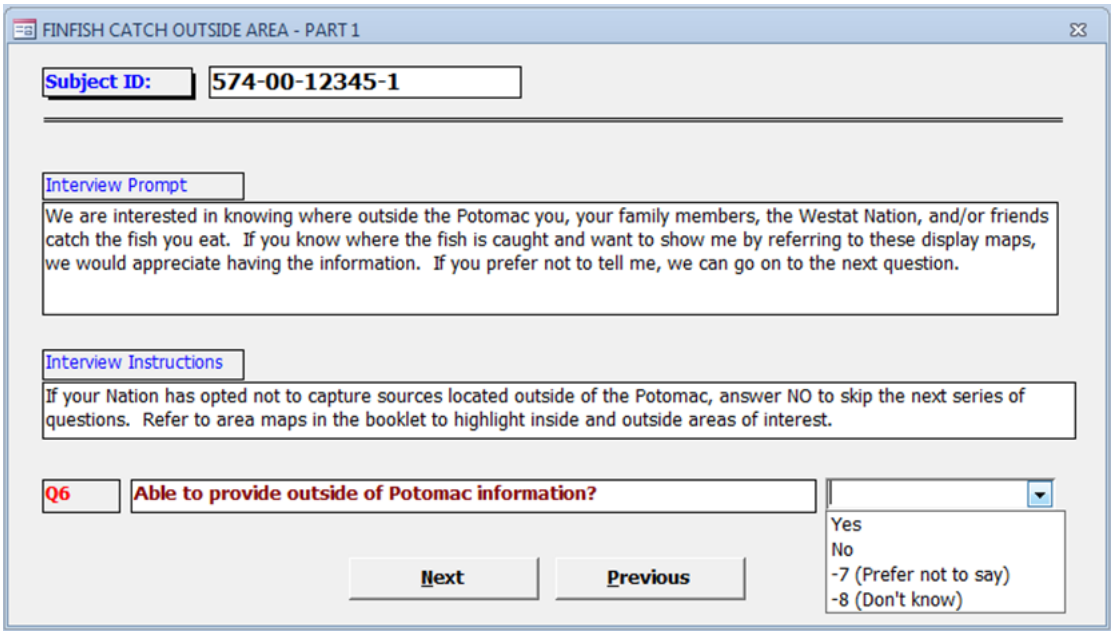


Figure 11. Interviewer should answer "No" to avoid prompting Respondent for seafood sources located outside the tribal area of interest.

Shellfish In this section, the Respondent is asked about shellfish consumption rates and amounts, shellfish sources (e.g. grocery store, restaurant, and harvest area), and parts of the shellfish eaten. This section is divided into multiple subsections, one for each of the group of shellfish defined by the Supervisor. After providing portion and frequency values for a specific shellfish, the Respondent is asked what percent of the time specified parts of that shellfish are eaten. As in the finfish groups, the Respondent is asked to consider all consumed shellfish when reporting sources. If the shellfish was harvested in or outside the area of interest, the Respondent is prompted to select locations from a pre-defined list for each shellfish item reported consumed. Again, some studies may choose not to collect information about seafood sources located outside the main area of interest. To skip questions that ask the respondent for outside locations answer "No" to this question (see Figure 9).

Other Seafood In this section, the Respondent is asked to report the consumption of other seafood as defined by the Supervisor. The Respondent is also given an opportunity to provide names of other seafood, not previously asked, and report on consumption.

Child In this section, Respondents are asked to report on children living in their households and seafood consumption patterns for each reported child under six years of age. If the study is not interested in fish consumption of children, do not ask respondent, just select "No" to Q31, "Are there any children under six years old living in your household?". Any household member can answer questions about a child, including a parent or grandparent. However, if any other household member has participated in the CAPI survey, the current Respondent should not answer these questions. Skip these questions by answering "Yes" to the prompt shown in Figure 12.

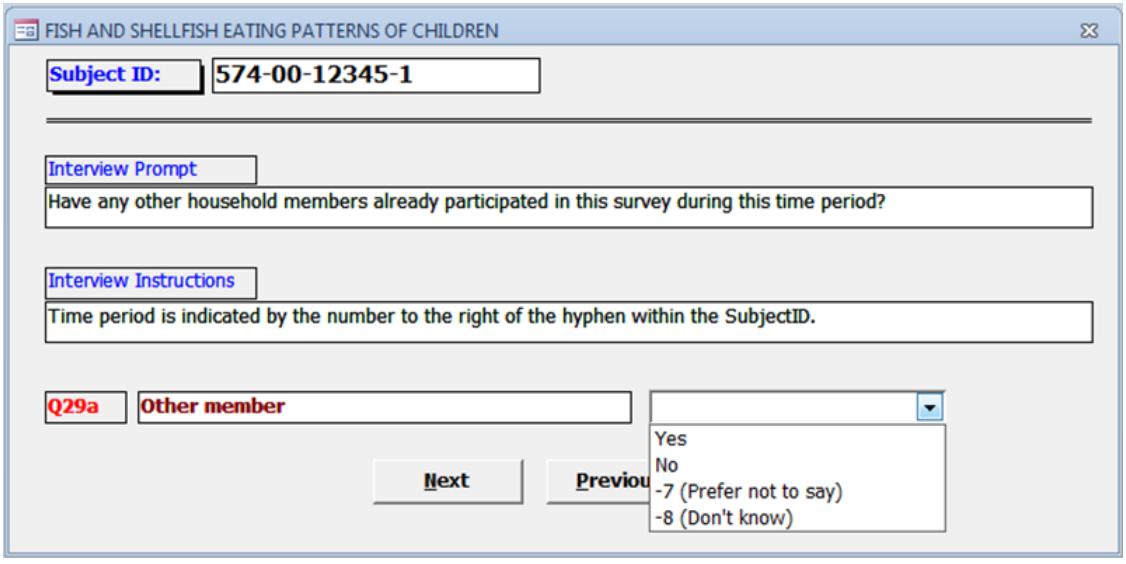


Figure 12. Interviewer should answer "Yes" to skip Children section if other household members previously responded to the CAPI Survey

For Finfish, the child questions have been configured to be asked at the group level and not at a specific species level. For each finfish group reported consumed, the Respondent is asked to report on the percentage of parts consumed. Shellfish group specific items, such as butter clams and horse clams, are then asked about, as are the percent of specific parts eaten for each item.

General and Special Events Questions In this section, the Respondent is prompted for general information on seafood consumption, demographic information, and consumption of certain seafood items at special events and gatherings. Also, in this section, you, the Interviewer, record that the Respondent signed the Participation Verification Form as well as the interview stop time and duration.

Interviewer Notes In this section, the Interviewer records any comments about how well the interview went and rates the reliability of the answers.

Starting the CAPI Application

During installation an icon was copied to your desktop.

Figure 13. Desktop icon

Double click on this icon to start the software. The Interviewer will be prompted for his/her name that will be used by the software to create a personalized introduction.

Main Menu Functions The main menu has four items: Exit, Data Entry, Export, and About.

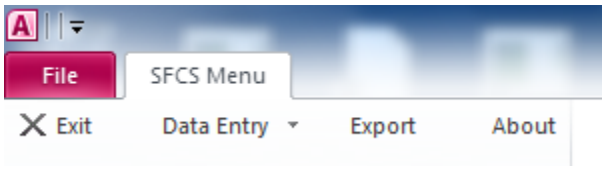


Figure 14. Main menu

Exit To exit the application, you must choose the Exit item under the File tab or under the SFCS Menu tab. The Windows close button that appears in the top

right hand corner of the screen has been disabled to prevent abnormal shutdown that might affect the integrity of the database.

Data Entry When this item is selected, three sub-menu items will appear: *New Interview*, *Edit Interview,* and *Delete Interview*. Select *New Interview* when starting a new interview. The software will assign a unique identifier to each subject's interview. Choose *Edit Interview* to open an interview that was previously completed or started. You will be prompted to provide the subject interview ID. Choose *Delete Interview* to delete an existing interview and remove all interview data from all data tables.

Export This function will prepare the database on your desktop for transfer to the Supervisor's PC where it will be integrated with CAPI data collected by other Interviewers. You will only be allowed to export your data once. It should be done ONLY when you have completed all interviews.

About The selection of this item causes a window to open that reports the version of the software. Hit the OK button to return to the main menu.

To begin an interview, select the Data Entry Tab and select "New Interview".

Screen Layout

Most of the CAPI screens are designed to ask one question per screen.

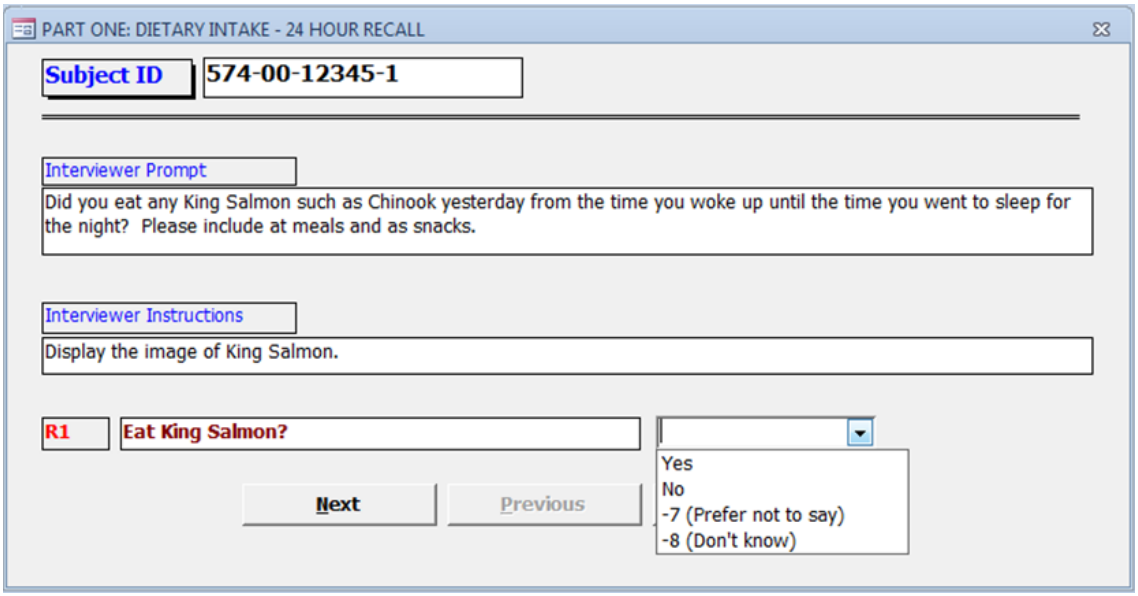


Figure 15. Typical screen layout

Below the Subject ID, there is an Interviewer Prompt. This is the text the Interviewer should read to the Respondent. It should be read exactly as it appears.

Next on the screen, Interviewer Instructions appear. This text should **not** be read to the Respondent but should be read silently by the Interviewer. This text should help the Interviewer better guide the Respondent and manage the interview. If there are no relevant instructions, the word *None* will appear in the box.

Below the Interviewer Instructions text box, the field name, a brief field description, and finally, the response field appear. The cursor will be focused on the response field where you will key the Respondent's answer. If there is a predefined codelist, a downward facing arrowhead will appear to the right of the field. Clicking on the arrowhead will display the relevant codelist. For numeric fields and non-codelist fields, the Interviewer must type in a value or text response.

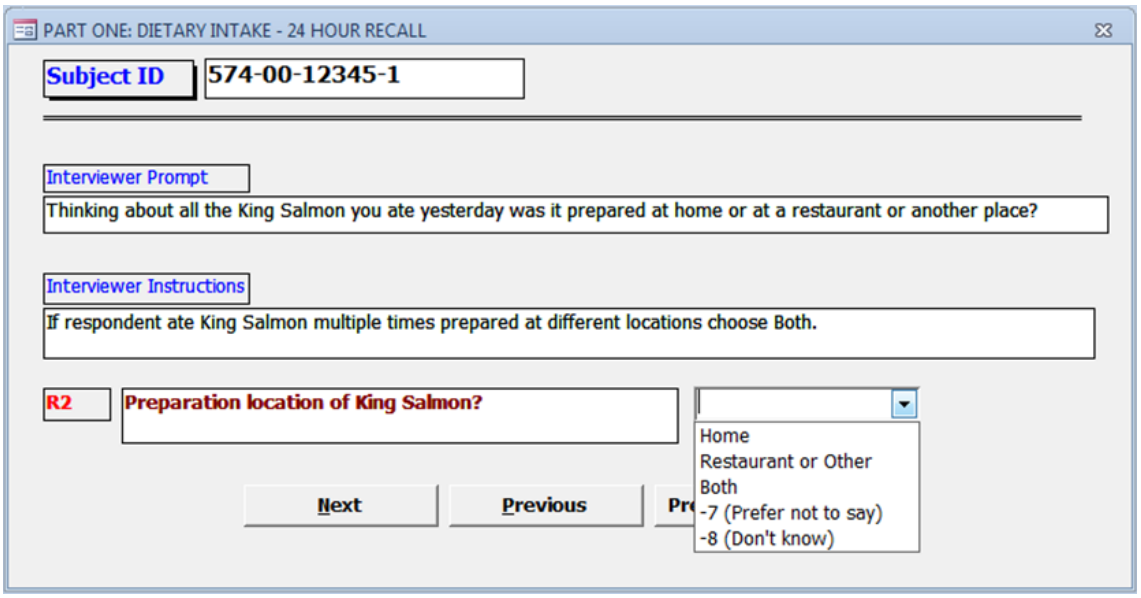


Figure 16. Field with a codelist

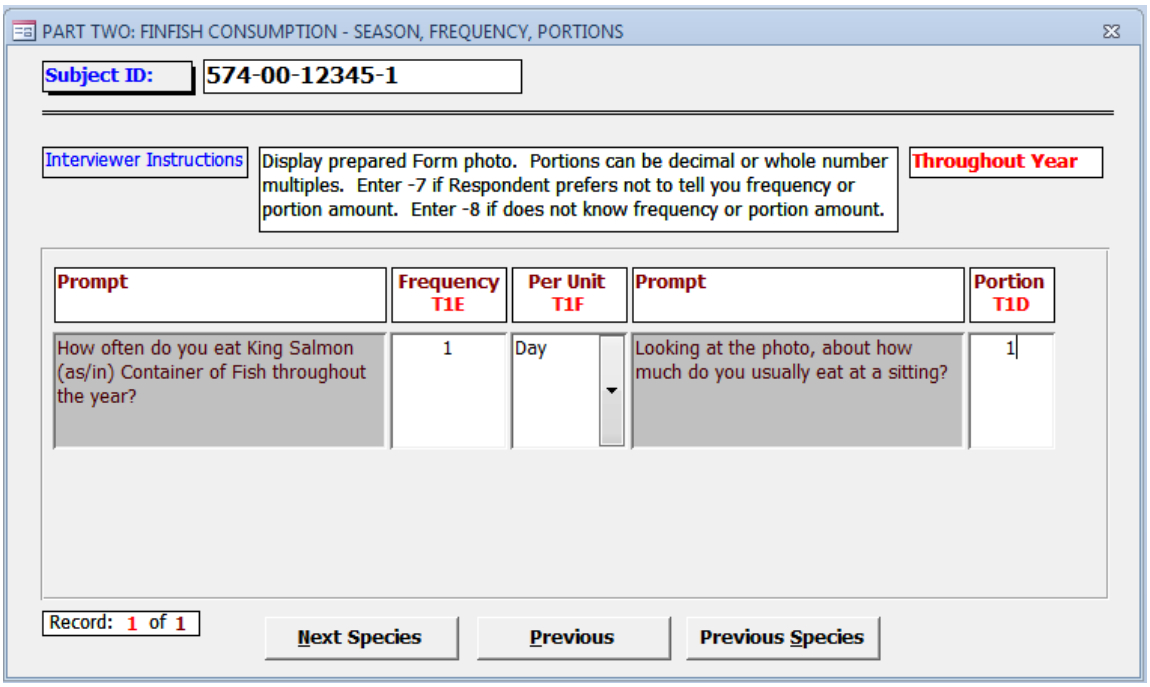


Figure 17. Numeric field

At the bottom of the screen, the active control buttons are displayed. These buttons will allow the Interviewer to navigate within the application.

There are a few screens in the CAPI that contain two or more fields that logically belong together. For example, frequency and unit,

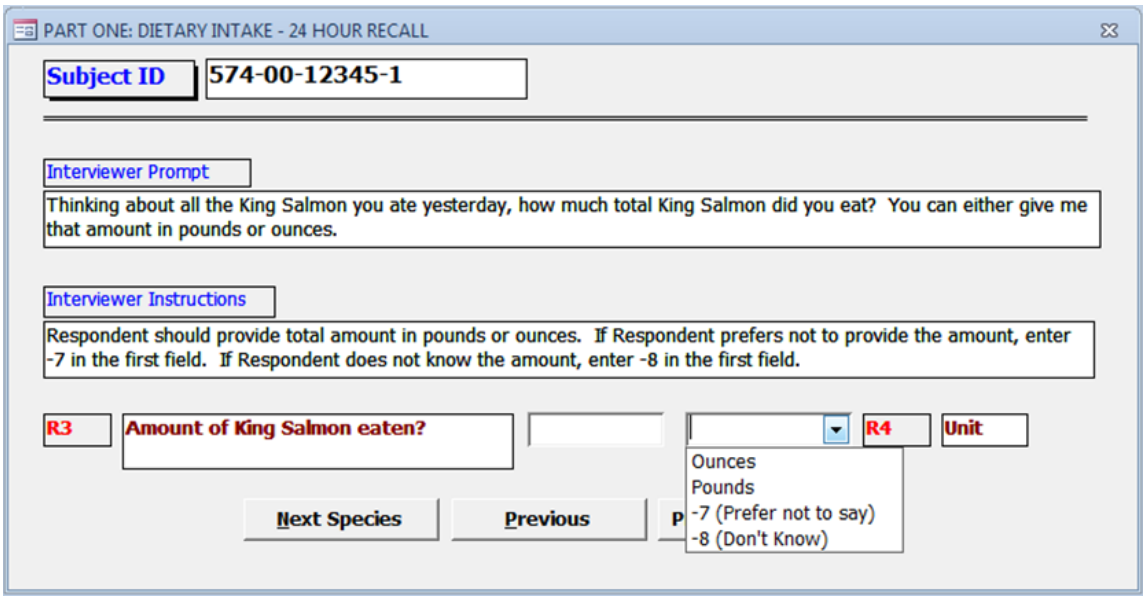


Figure 18. Screen with two fields

height, feet, and inches,

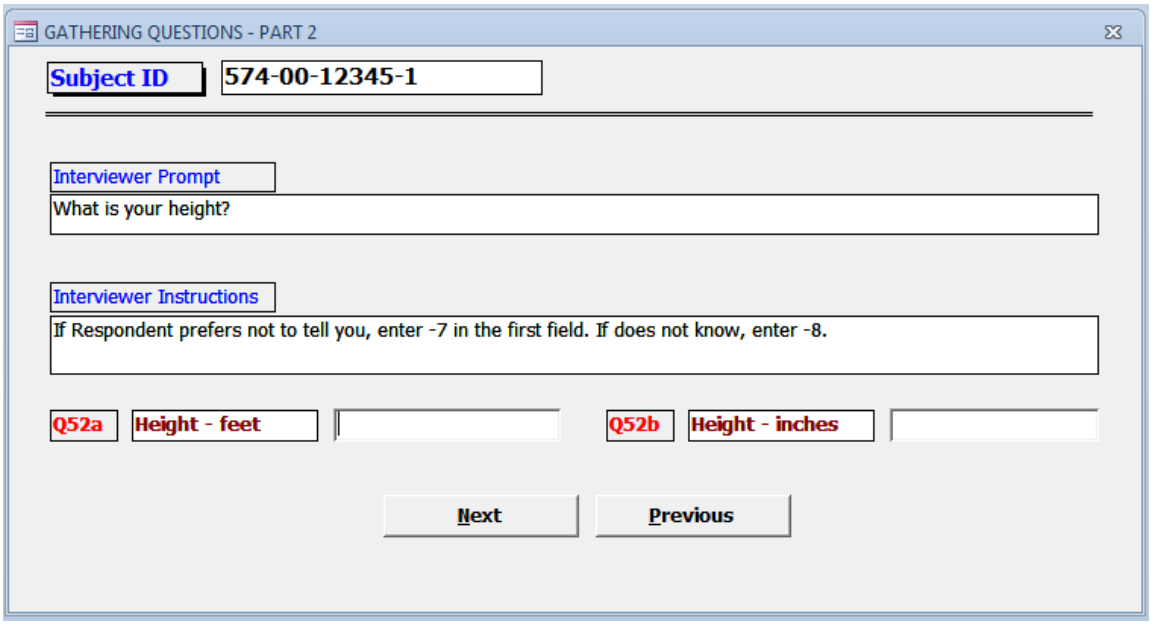


Figure 19. Another screen with two fields

and percents that must add to 100:

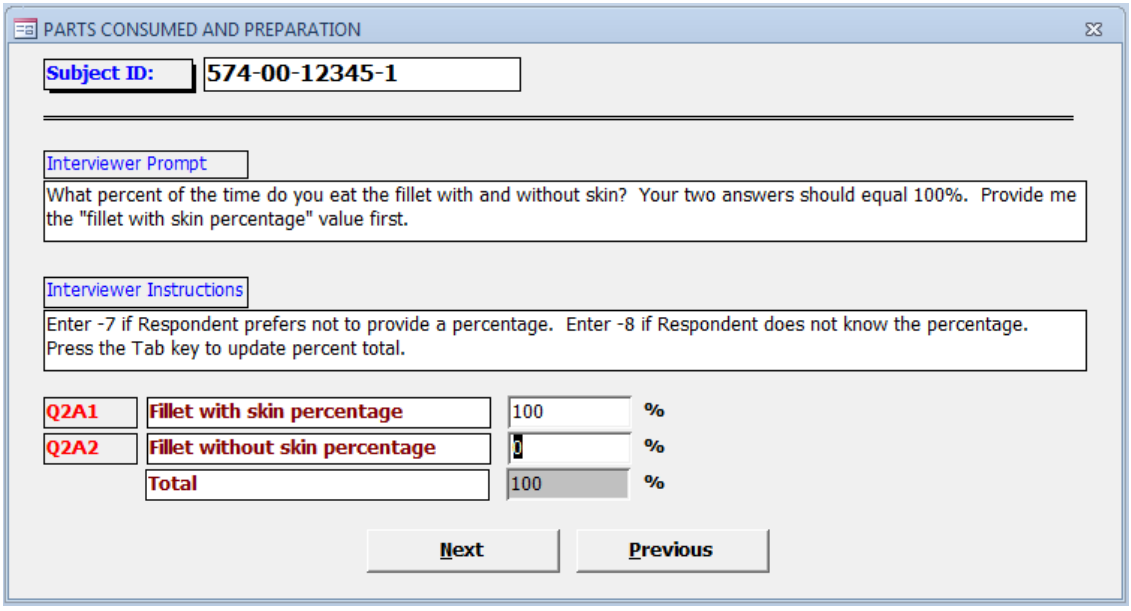


Figure 20. Screen with multiple fields

Comments

The Interviewer can record comments or notes at anytime during an interview. When the **F3** key is hit, the Comments Window is opened.

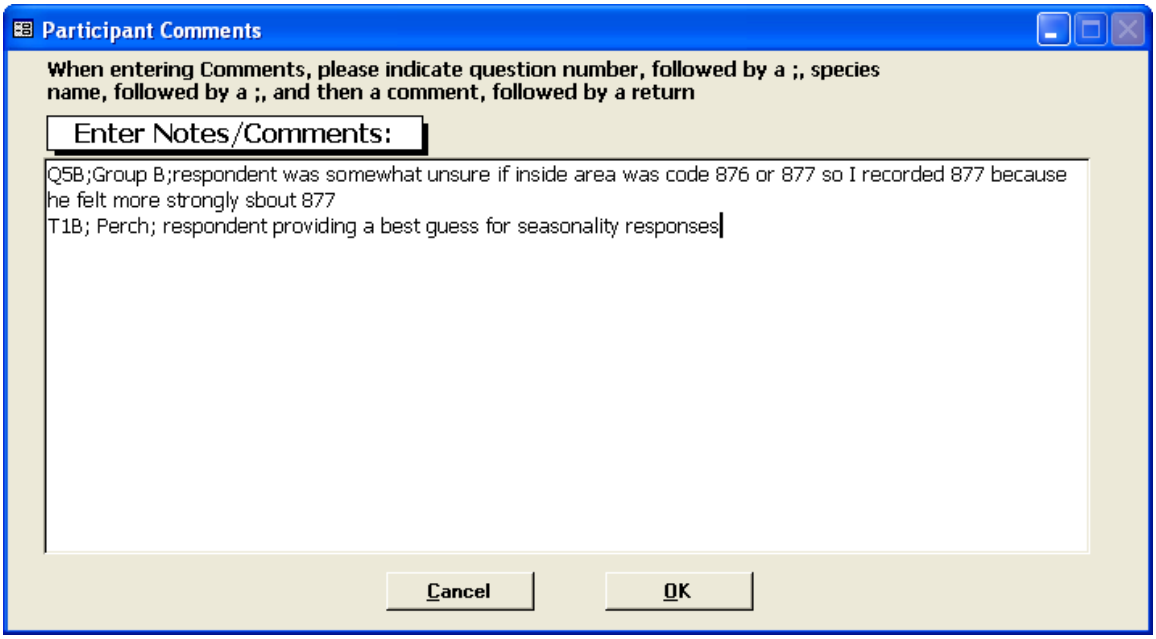


Figure 21. Comments window

As stated in the top of the Comments window, the Interviewer should record comments in a standard format to facilitate any future analyses. In the above example, the Interviewer has two comments thus far indicating for "Group B" finfish, the Respondent was somewhat unsure about a catch area but felt strong enough to provide an answer and for Perch was providing a best guess for seasonality.

The Interviewer can minimize the window by hitting the button  in the top right hand corner of the Comments window. The Comments window will then appear minimized in the bottom left hand corner of the Access screen. To maximize this window, the Interviewer can click on the window's restore control or hit **F3** again. Any additions and changes have not yet been saved.

The Interviewer can also close the Comments window and save the changes by hitting the **OK** button. Again, **F3** will open the window.

The Interviewer can hit the Cancel button to not save the most recently entered comments, that is, since the **OK** button was hit last.

Missing Values A value is required for most every CAPI question field, otherwise the software will not allow the interview to proceed. For most of the questions, the Interviewer is allowed to record "missing" responses. A value of **-7** should be

recorded in a non-codelist or numeric field when the Respondent refuses to answer. A value of **-8** should be recorded in a non-codelist or numeric field when the Respondent does not know the answer. Codelists accessed for fields with the downward facing arrowhead, contain these two missing response codes.

Edits and Skips When the Interviewer attempts to move off a screen, pre-defined edit checks are run on the entered value(s). If the value falls outside a prescribed range or is inconsistent, given another response(s), the Interviewer will be presented with a message window that provides information so that the value can be verified.

If there is a range or logic error, the message will tell the Interviewer why the value is suspect.

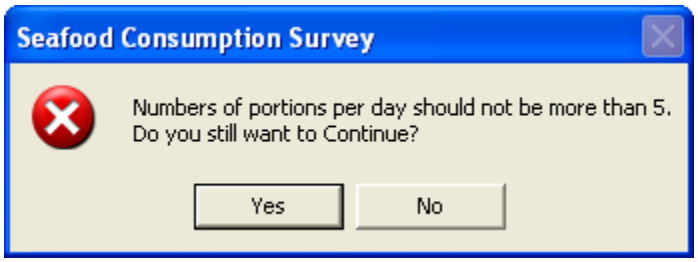


Figure 22. Example of warning message

These types of errors can be ignored and over-ridden by the Interviewer by hitting the **Yes** button to continue. Or the Interviewer can choose to remain on the field if the **No** button is selected, so as to review the response and make any necessary corrections.

If a value that was provided or changed conflicts with other information, the message will inform the Interviewer of the conflict and, if appropriate, provide a warning that the application will clear out any now irrelevant data.

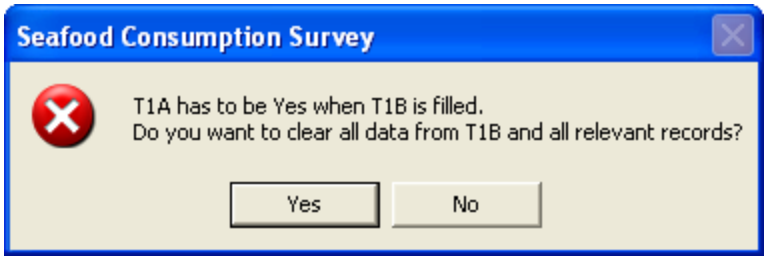


Figure 23. Example of warning message

Some questions will be skipped if a previous response or lead-in type question is answered a certain way. For example, if the Respondent is unable to answer questions about her child's diet, all the dietary questions will be skipped and the interview will continue with the general questions.

Navigation There are two basic ways to navigate within an ongoing interview: the control buttons at the bottom of each screen and the *GoTo* Menu item at the top of the screen.

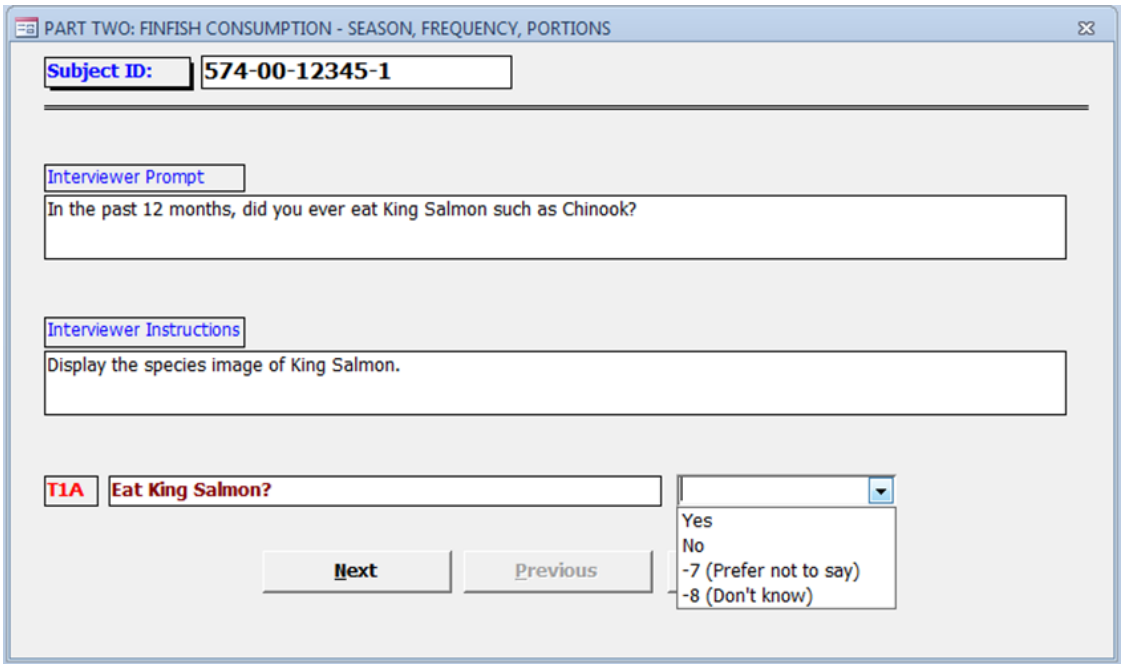


Figure 24. Typical screen view

The buttons at the bottom of the screen will vary depending on the CAPI section and active question. The buttons allow the Interviewer to navigate within one of the CAPI sections. If the Interviewer wants to move to a previously completed section, the Interviewer must use the *GoTo* menu item. Buttons will be grayed out and rendered inactive if not appropriate for the species and/or question in focus.

The **Next button** will advance the Interviewer to the next screen in the interview sequence. Any prescribed edits will be run before the next screen is displayed. Prescribed skipped patterns will be followed. For example, if the Respondent answered that his consumption pattern had not changed over the past 20 years, the **Next** button will skip to a question about frequency of special events attendance. If the Respondent had answered that his diet had changed, the **Next** button would advance the Interviewer to the question about how it has changed.

When on the last question for a food item, the **Next** button will be inactive.

The **Previous** button will take the Interviewer to the previous screen or lead-in question that controlled a series of questions.

For example, if the Interviewer was asking about out of season consumption of King or Chinook Salmon

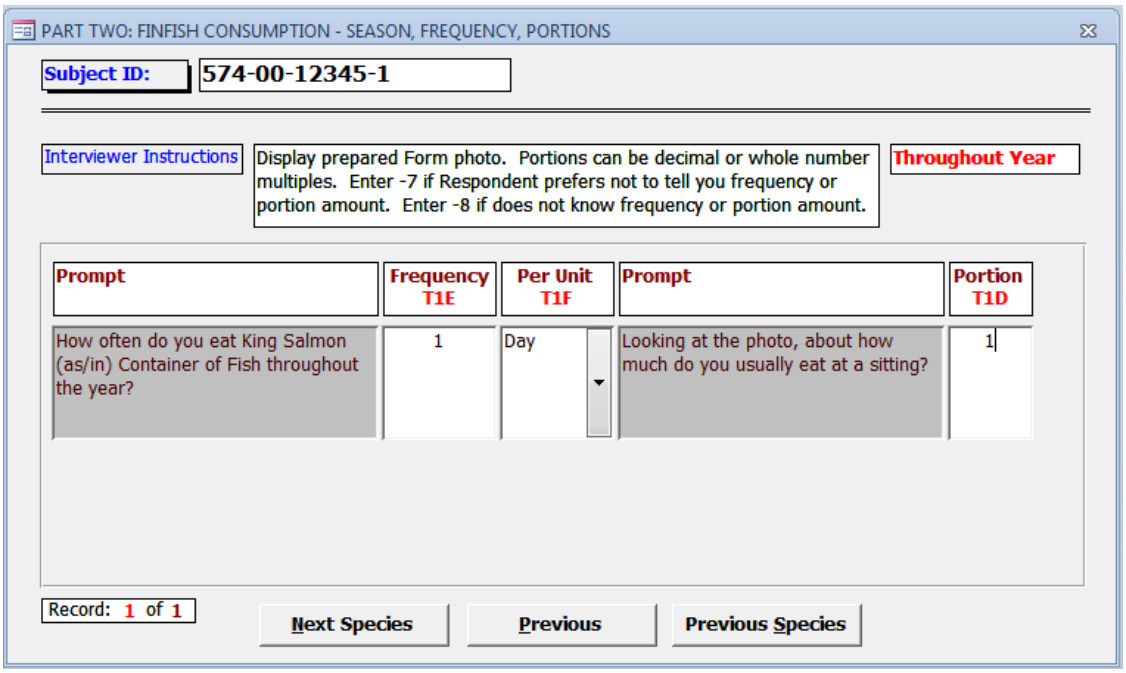


Figure 25. Example of active Previous button

and hit the **Previous** button, the Interviewer would be taken back to questions about in season consumption.

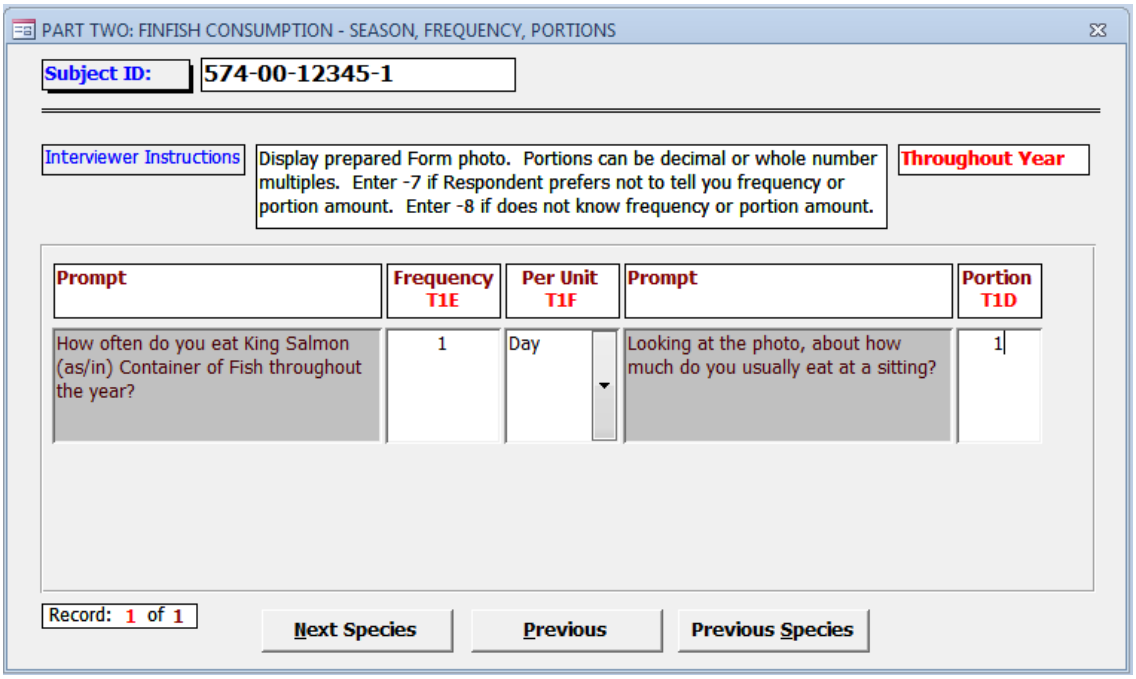


Figure 26. Preceding screen to Figure 23

If the active question is the first in a section or series of questions, the **Previous** button will be gray and inactive. In this example, the question *Do you eat King Salmon such as Chinook?* is the first in the series of frequency questions for this food item.

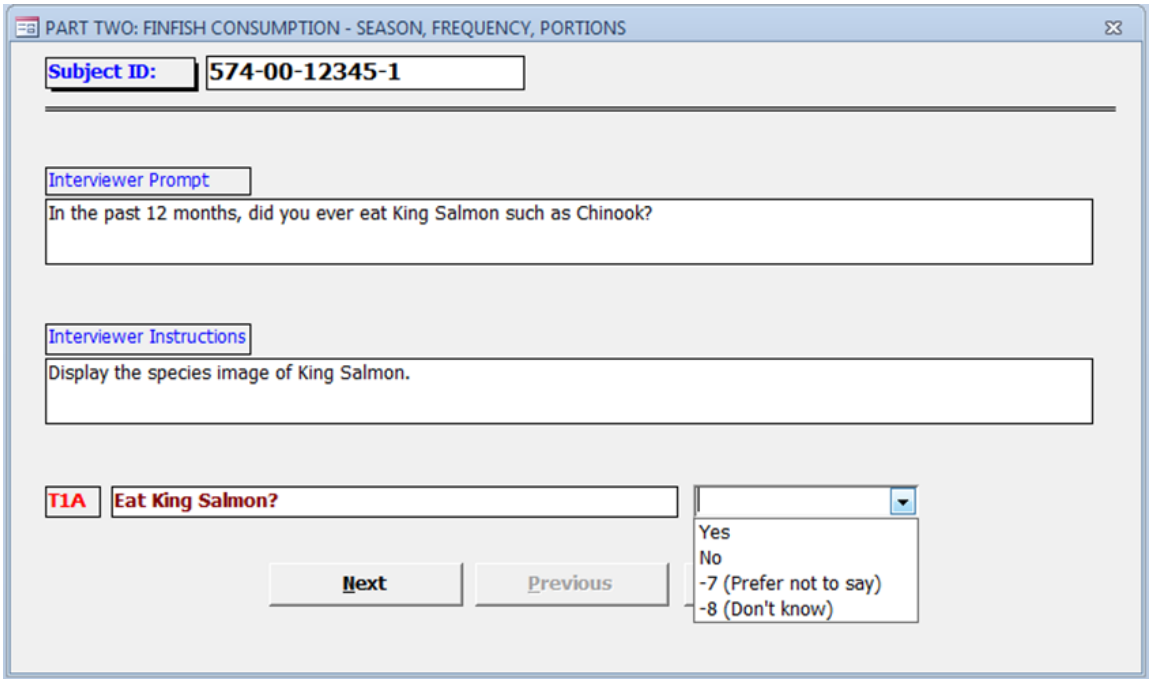


Figure 27. First question for cod

The **Next Species** button will be active on the last question of a series of questions asked for each species item. Selecting this button will advance the Interviewer to the next species item in that group. If the current species is the last item in the group, the interview will proceed to the next question type.

The **Previous Species** button will be active for all but the first food item in a group. At any point during a question, selecting this button will return the user to the first question in the frequency series for the most recently asked food item.

GoTo The *GoTo* menu item is accessible during an interview if it becomes necessary to return to a previously completed section. When this item is selected, a window of completed sections will be displayed so that the Interviewer can choose a section to move to.

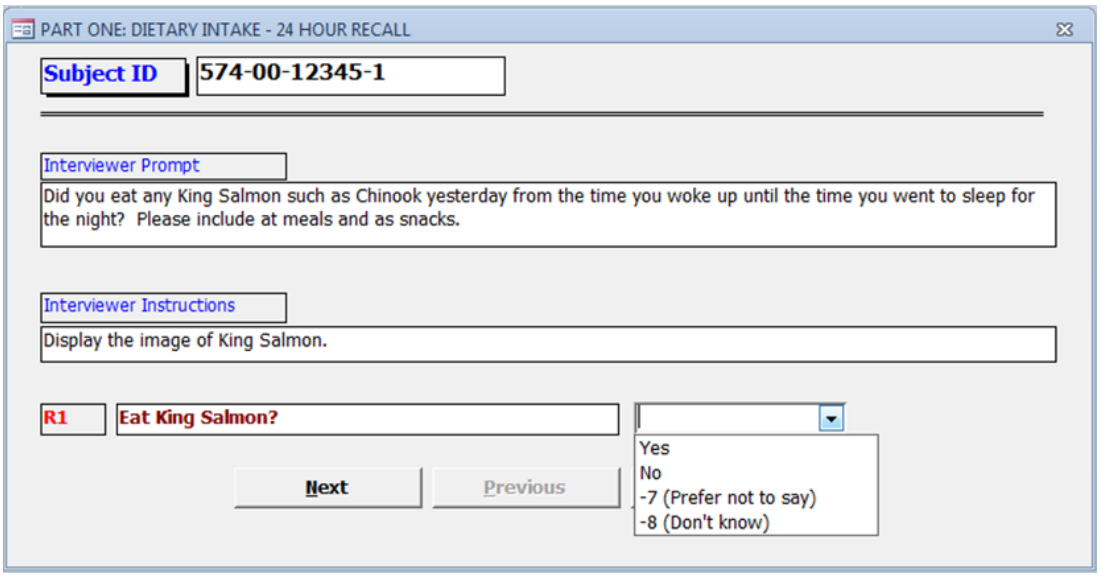


Figure 28. GoTo menu item

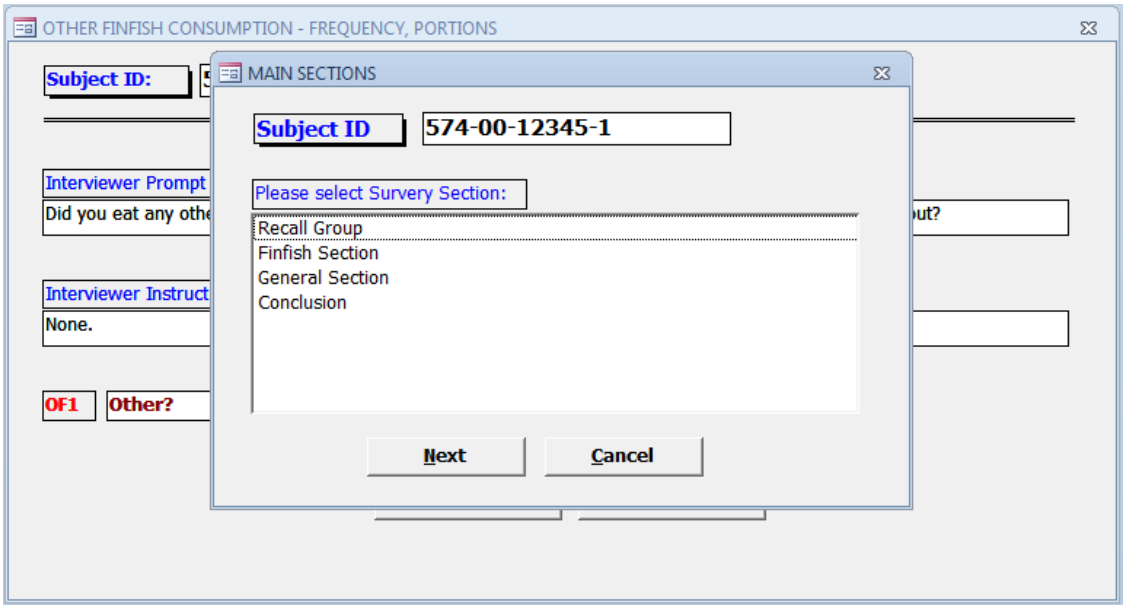


Figure 29. GoTo brings up previously completed sections

Break If it becomes necessary to stop an interview in mid-stream, the Interviewer can select the Break menu item at the top of the screen. This will cleanly shut down the interview and return the Interviewer to the main menu. To re-enter an interview that has been "broken off" or completed, select *Edit Interview* under the Data Entry menu item and provide the subject interview ID. A window of previously touched sections will appear. The Interviewer can then select where to begin again.

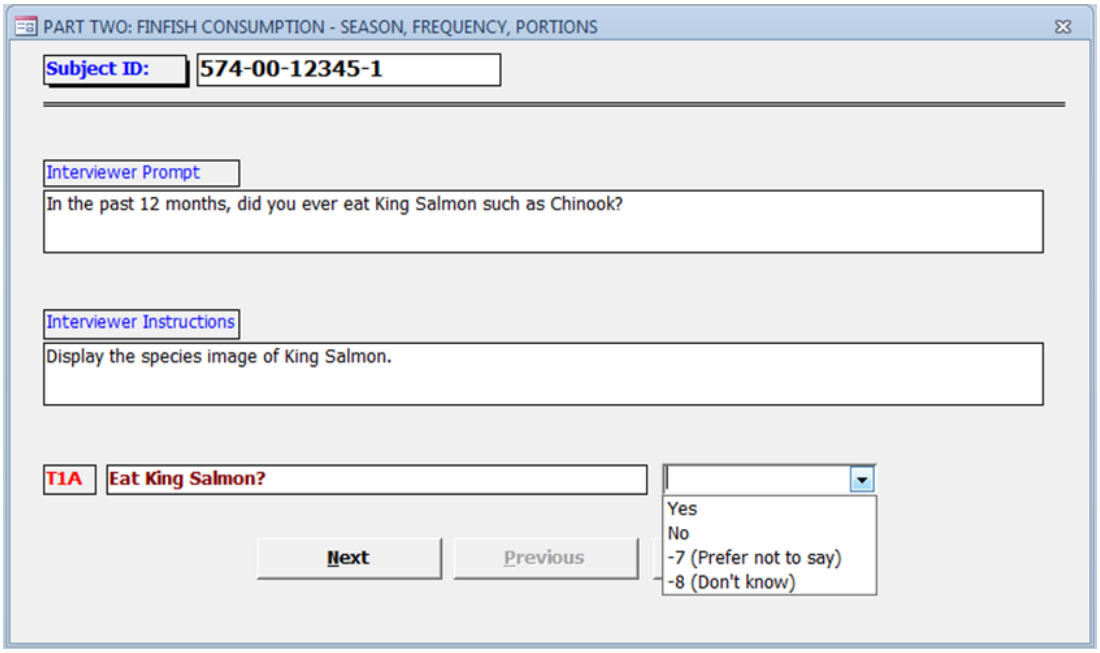


Figure 30. Break menu item returns you to main menu

Export: Once the interviewing phase is over, you will need to send your data to the Supervisor for integration and analysis. You should have ready and inserted into a port, a USB device (also known as a thumb-drive or flash-drive) before you choose this function. The **Export** function will compress the database containing the interview data and copy the "zipped" file to the USB drive.

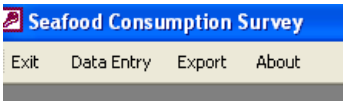


Figure 31. Export menu item

You will be prompted to provide the location where the files will be written.

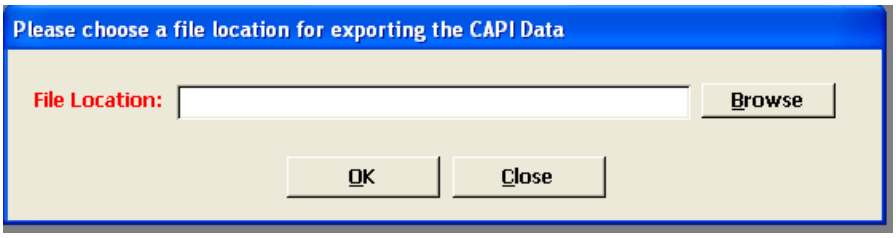


Figure 32. Prompt to write to USB drive

Choose the **Browse** button to help navigate to the USB port. Specify the drive letter that corresponds to the USB device which in Figure 31 is drive "**E:**".

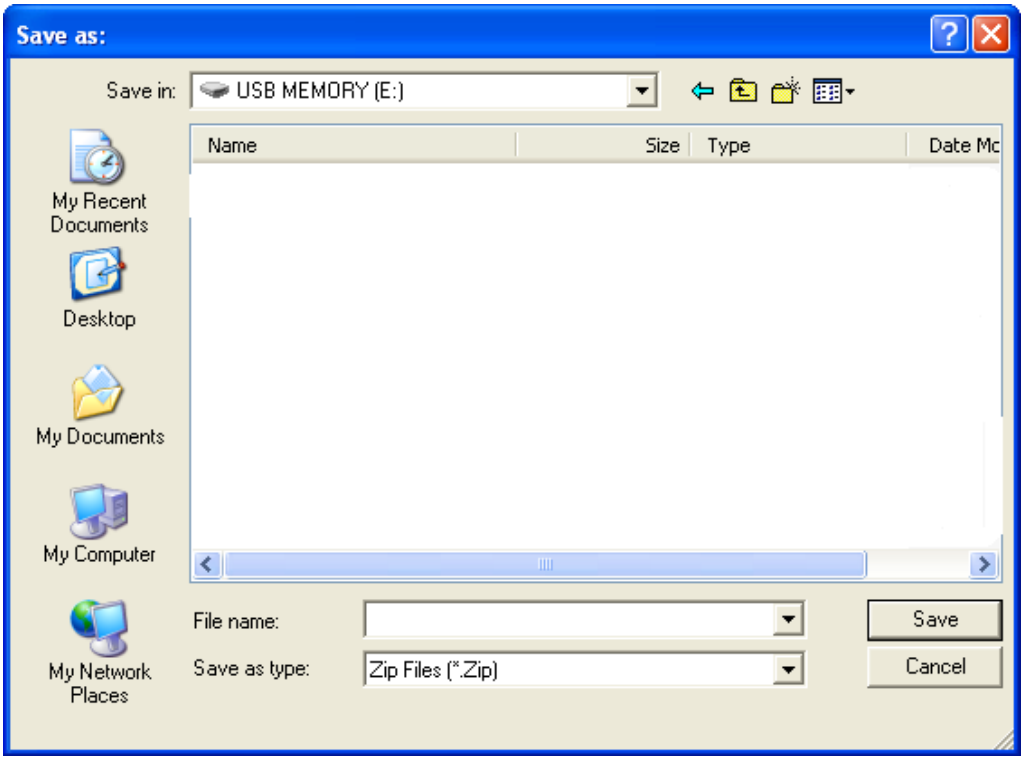


Figure 33. Specify the filename as your site name

Choose **Save** to export the interview data in compressed format to the USB device. The Export function will not delete any files already on the device. Send the device to the Supervisor for processing.

You will only be allowed to export once.