SECTION 9

DATA ANALYSIS AND REPORTING

This section provides guidance on (1) analysis of laboratory data for both screening and intensive studies that should be included in state data reports, (2) data reporting requirements for both state-conducted screening and intensive studies, and (3) data reporting requirements for a national data repository for state-collected fish tissue data housed within the National Listing of Fish and Wildlife Advisories (NLFWA) database.

All data analysis and reporting procedures should be documented fully as part of the Work/QA Project Plan for each study, prior to initiating the study (see Appendix I). All routine data analysis and reporting procedures should be described in standard operating procedures. In particular, the procedures to be used to determine if the concentration of a target analyte in fish or shellfish tissue differs significantly from the selected screening value must be clearly documented.

9.1 DATA ANALYSIS

9.1.1 Screening Studies

The primary objective of **Tier 1** screening studies is to assist states in identifying potentially contaminated harvest areas where further investigation of fish and shellfish contamination may be warranted. The criteria used to determine whether the measured target analyte concentration in a fish or shellfish tissue composite sample is different from the SV (greater than or less than) should be clearly documented. If a reported target analyte concentration exceeds the SV in the screening study, a state should initiate a **Tier 2**, **Phase I**, intensive study (see Section 6.1.2.1) to verify the level of contamination in the target species. Because of resource limitations, some states may choose to conduct a risk assessment using screening study data; however, this approach is not recommended because a valid statistical analysis cannot be performed on a single composite sample. If a reported analyte concentration is close to the SV but does not exceed the SV, the state should reexamine historic data on water, sediment, and fish tissue contamination at the site and evaluate data on laboratory performance. If these data indicate that further examination of the site is warranted, the state should initiate a **Tier 2**, **Phase I**, intensive study to verify the magnitude of the contamination.

Because replicate composite samples are not required as part of a screening study, estimating the variability of the composite target analyte concentration at any site is precluded. The following procedure is recommended for use by states for analysis of the individual target analyte concentration for each composite sample from reported laboratory data (see Section 8.3.3.3)

- A datum reported below the method detection limit, including a datum reported as not detected (i.e., ND, no observed response) should be assigned a value of one-half the MDL or zero.
- A datum reported between the MDL and the method quantitation limit should be assigned a value of the MDL plus one-half the difference between the MQL and the MDL.
- A datum reported at or above the MQL should be used as reported.

This approach is similar to that published in 40 CFR Parts 122, 123, 131, and 132—Proposed Water Quality Guidance for the Great Lakes System.

If resources permit and replicate composite samples are collected at a suspected site of contamination, then a state may conduct a statistical analysis of differences between the mean target analyte concentration and the SV, as described in Section 9.1.2.

9.1.2 Intensive Studies

The primary objectives of **Tier 2** intensive studies are to confirm the findings of the screening study by assessing the magnitude and geographic extent of the contamination in various size classes of selected target species. The EPA Office of Water recommends that states collect replicate composite samples of three size classes of each target species in the study area to verify whether the mean target analyte concentration of replicate composite samples for any size class exceeds the SV for any target analyte identified in the screening study. The statistical approach for this comparison is described in Section 6.1.2.7.

The following procedure is recommended for use by states in calculating the mean arithmetic target analyte concentration from reported laboratory data (see Section 8.3.3.3.3).

- Data reported below the MDL, including data reported as not detected (i.e., ND, no observed response) should be assigned a value of one-half the MDL.
- Data reported between the MDL and the MQL should be assigned a value of the MDL plus one-half the difference between the MQL and the MDL.
- Data reported at or above the MQL should be used as reported.

This approach is similar to that published in 40 CFR Parts 122, 123, 131, and 132—Proposed Water Quality Guidance for the Great Lakes System.

Secondary objectives that may be assessed as part of **Tier 2** intensive studies

can include defining the geographical region where fish contaminant concentrations exceed screening values; identifying geographical distribution of contaminant concentrations; and, in conjunction with historical data or future data collection, assessing changes in fish contaminant concentrations over time. The statistical considerations involved in comparing fish contaminant levels measured at different locations or times are discussed in Appendix N.

State staff should consult a statistician in interpreting intensive study tissue residue results to determine the need for additional monitoring, risk assessment, and issuance of a fish or shellfish consumption advisory. Additional information on risk assessment, risk management, and risk communication procedures will be provided in later volumes in this guidance series (see Section 1.4).

9.2 DATA REPORTING

9.2.1 State Data Reports

State data reports should be prepared by the fish contaminant monitoring program manager responsible for designing the screening and intensive studies. Summaries of **Tier 1** screening study data should be prepared for each target species sampled at each screening site. For **Tier 2** intensive studies (**Phase I** and **Phase II**), data reports should be prepared for each target species (by size class, as appropriate) at each sampling site within the waterbody under investigation (see Section 6.1.2). Screening and intensive study data reports should include, at a minimum, the information shown in Figure 9-1.

9.2.2 Reports to the National Fish Tissue Residue Data Repository (NFTRDR)

The EPA Office of Science and Technology within the Office of Water has established the NFTRDR, which is housed within the NLFWA database. This repository is a collection of fish and shellfish contaminant monitoring data gathered by various state, federal, and local agencies for advisory purposes. The objectives of the repository are to:

- Facilitate the exchange of fish and shellfish contaminant monitoring data nationally by improving the comparability and integrity of state data
- Encourage greater cooperation among regional and state fish advisory programs
- Assist states in their fish tissue data collection efforts by providing ongoing technical assistance.

The NLFWA database now contains a facility for storing fish tissue residue data as well as for documenting and mapping active and rescinded fish consumption advisories. Since 1996, a stand-alone version of the NFLWA database has been available for Internet downloads. Internet WEB-based tools have recently been developed to support queries and interactive mapping of both the general advisory information as well as fish tissue residue data. Internet-based tools are also being

- Study identification (e.g., project number, title, and study type)
- Program manager
- Sampling site name
- Latitude (decimal degrees preferred)
- Longitude (decimal degrees preferred)
- Type of waterbody (lake, river, estuary, etc.)
- Name of waterbody
- Sampling date (e.g., YYYYMMDD, Year 2000 compliant format)
- Sampling time (e.g., HH, MM in a 24-h format)
- Sampling gear type used (e.g., dredge, seine, trawl, gill net)
- Sampling depth (feet or meters)
- Standard common name of target species (preferably name sanctional by American Fisheries Society for inland waters or by NOAA for coastal waters)
- Composite sample numbers
- Number of individuals in each composite sample
- Number of replicate composite samples
- Predominant characteristics of specimens used in each composite sample
  - Predominant life stage of individuals in composite
  - Predominant sex of individuals in composite (if applicable)
  - Average age of individuals in composite (if applicable)
  - Average body length (cm)
  - Description of edible portion (tissue type)
- Analytical methods used (including method for lipid analysis)
- Method detection and quantitation limits for each target analyte analyzed
- Sample cleanup procedures (e.g., additional steps taken to further purify the sample extracts or digestates)
- Data qualifiers (e.g., additional qualifying information about the measurement)
- Percent lipid (wet weight basis) in each composite sample
- For each target analyte in each composite sample:
  - Total wet weight of composite sample (g) used in analysis
  - Measured concentration (wet weight basis) as reported by the laboratory (see Section 8.3.3.3.3)
  - Units of measurement for target analyte concentration (e.g., ppm)
  - Evaluation of laboratory performance (i.e., description of all QA and QC samples associated with the sample(s) and results of all QA and QC analyses)
- In screening studies with only one composite sample for each target species, the state should provide for each target analyte a comparison of reported concentration with selected SV (see Section 4 tables) and indication of whether SV was exceeded for recreational or subsistence fishers (see Section 9.1.1).
- In intensive studies, for each target analyte in each set of replicate composite samples, the state should provide
  - Range of target analyte concentrations for each set of replicate composite samples
  - Mean (arithmetic) target analyte concentration for each set of replicate composite samples (see Section 9.1.2)
  - Standard deviation of mean target analyte concentration
  - Comparison of target analyte arithmetic mean concentration with selected SV (see Section 5) and indication of whether SV was exceeded for recreational or subsistence fishers

Figure 9-1. Recommended data reporting requirements for screening and intensive studies.

developed as a way for state agencies to add fish advisory and contaminant monitoring data to the NLFWA database and may be developed to perform some types of standard data analysis on the fish tissue residue data.

EPA has recently developed an Internet-based data entry facility for the NLFWA using some of the data elements included in Figure 9-1. This Internet-based data entry facility is housed within the EPA's NLFWA database and allows states to archive fish advisory information as well as fish tissue residue data generated through their fish contaminant monitoring programs. States may prepare their own data tables and arrange to transfer these to EPA to be formatted and reviewed before entry into the repository. The information in the NFTRDR can be organized into three different tables (STATIONS, SAMPLES, and RESULTS tables) using such readily available PC relational database packages as ACCESS (Figure 9-2). If states submit their monitoring data in other file formats (e.g., spreadsheet files or ASCII files exported from other in-house database systems), a short data dictionary (metadata) file should be included (ASCII, Wordperfect, or WORD format) clearly documenting the meaning of all data fields and any codes, abbreviations, or measurement units used in the files.

State, regional, and local agency staff may obtain further information on the new Internet WEB-based database EPA now has available by contacting:

U.S. Environmental Protection Agency
Office of Science and Technology
National Fish and Wildlife Contamination Program-4305
1200 Pennsylvania Avenue, NW
Washington, DC 20460
PHONE: 202-260-7301
FAX: 202-260-9830

Jeffrey D. Bigler
U.S. Environmental Protection Agency-4305
1200 Pennsylvania Avenue, NW
Washington, DC 20460
PHONE: 202-260-1305
E-MAIL: bigler.jeff@epa.gov

Fish Tissue Chemical Residue Data Tables: STATIONS, SAMPLES and RESULTS

The **STATIONS** table includes basic locational data.

Field name	Field description
STATION_ID	Waterbody, Station or Monitoring Site Identifier. This field becomes a database key field. Each record must have a unique STATION_ID.
STATE	State 2-character postal code abbreviation.
WATERBODY (or SITENAME)	A short caption to identify the waterbody or sampling station.
LOCATION	Additional descriptive information on the waterbody or station location.
ADVNUM	If the waterbody or site is associated with an advisory (active or rescinded), include the number assigned to this advisory in the current National Listing of Fish and Wildlife Advisories (NLFWA) database.
COUNTY	County name.
LAT	Station latitude. A format in decimal degrees is preferred.
LNG	Station longitude. A format in decimal degrees is preferred.

The **SAMPLES** table includes data on the type of tissue sample collected.

Field name	Field description
SAMPLE_ID	An identifier to each specific fish tissue sample from a waterbody or station. This is used as a database key, so each record must have a unique SAMPLE_ID
STATION_ID	Waterbody, Station or Monitoring Site Identifier as defined in the STATIONS table.
SAMPLE_DATE	The date the sample was collected in the field. Give date in a Year 2000 compliant format (YYYYMMDD).
FISH_SPECIES	Fish species names. Standard English common names as established by the American Fisheries Society for inland waters or NOAA for coastal water are preferred.
SAMPLE_TYPE	How the sample was prepared (e.g., fillet with skin-on or skin-off, whole fish). In the NUMBER_OF_FISH field below, multiple fish in a sample indicate a composite sample.
LENGTH	The length of the sample fish. For composites, an average length should be given.
LENGTH_UNIT	Length units of fish (cm or inches)
WEIGHT	Specimen or composite weight used for residue analysis.
WEIGHT_UNIT	Weight units (usually in grams).
LIPID	Percent extractable lipids.
NUMBER_OF_FISH	Number of fish (specimens) in sample. Number greater than a value of 1 indicates a composite sample.

The **RESULTS** table includes chemical-specific tissue sample concentrations.

Field name	Field description
SAMPLE_ID	An identifier to each specific fish tissue sample from a waterbody or station. This is used as a database key, so each record must have a unique SAMPLE_ID
PARAMETER	Chemical name. File should specify all acronyms or abbreviations used.
DETECTION_INFO	A caption to document detection limit information (e.g., “less than detection limit”).
RESULT	A number representing the concentration of a chemical (or the detection limit).
RESULT_UNIT	Units associated with concentration (e.g., “ppm”).

Figure 9-2. Key information fields for the National Fish Tissue Residue Data Repository.