Science Advisory Board Consultation Document
Proposed Revisions to Aquatic Life Guidelines

Tissue-Based Criteria for "Bioaccumulative" Chemicals

Prepared by the
Tissue-based Criteria Subcommittee

August 2005

NOTICE

THIS DOCUMENT IS A PRELIMINARY DRAFT

It has been prepared for consultation with U.S. EPA's Science Advisory Board. It has not been formally released by the U.S. Environmental Protection Agency and should not at this stage be construed to represent Agency guidance or policy.

DISCLAIMER

This document is draft for Science Advisory Board consultation purposes only. It has been prepared by a technical working group as the basis for discussing scientific issues with the Science Advisory Board. This document has not undergone management or programmatic policy review. It has not been formally released by the U.S. Environmental Protection Agency and does therefore not constitute U.S. Environmental Protection Agency guidance or policy. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

Table of Contents

1 Introduction

1.1 Purpose and Scope of this Document

This document summarizes the current thinking of the U.S. Environmental Protection Agency (EPA)/U.S. Department of the Interior (DOI) Tissue-based Criteria Subcommittee (hereafter called the "Subcommittee") regarding approaches for revising EPA's Aquatic Life Criteria Guidelines to address risks from so called "bioaccumulative" chemicals. The Subcommittee drafted this document specifically to facilitate the 2005 Science Advisory Board (SAB) consultation on EPA's revision of its Aquatic Life Criteria Guidelines. Although nearly all chemicals bioaccumulate to some degree in aquatic organisms, we use the term "bioaccumulative" in this document to delineate chemicals which bioaccumulate extensively in aquatic food webs such that exposure from the diet becomes toxicologically important to relevant ecological receptors. Such chemicals generally persist in the aquatic environment, exhibit high hydrophobicity (e.g., log K_{ow} generally > 5), and are poorly metabolized by aquatic biota. The ecological receptors of primary concern for bioaccumulative chemicals in aquatic systems include aquatic life and aquatic-dependent wildlife (i.e., terrestrial wildlife that feed extensively on aquatic organisms).

We have organized this document into five main sections. First is an introduction that presents a brief history of EPA aquatic life criteria and exposes the need for a revised methodology that specifically addresses bioaccumulative chemicals. The second section provides the rationale and an overview of the proposed tissue-based criteria approach for both aquatic life and aquatic-dependent wildlife. The third and fourth sections present the salient features of tissue-based criteria for aquatic life and wildlife, respectively. The last section presents technical issues that the Subcommittee is seeking input from the SAB.

We think it is important to point out that this document represents "a work in progress" and that many issues and ideas have yet to be fully discussed or even explored by the Subcommittee. For example, the Subcommittee focused to date on national-level criteria but we recognize the need to address regional or site-specific criteria that account for regional or site-specific concerns. Further, the Subcommittee has not discussed in detail the application of population models for deriving aquatic life criteria for bioaccumulative chemicals, choosing instead to defer to the work of the water-based criteria subcommittee since most of the methodological issues will be the same. Finally, the Subcommittee clearly recognizes that the concept of tissue-based criteria can be appropriate for other types of chemicals (i.e., not just those where dietary exposure is important) and in particular where mixtures of chemicals with similar modes of action are of concern. To date, however, the Subcommittee has chosen to focus its efforts first on bioaccumulative chemicals due to concerns with the ability of existing Guidelines to adequately address risks from this group of chemicals.

The criteria process outlined in this document strives to make the best use of the data currently available. As this process evolves and guidance is developed for use by implementers in State, Tribal and local agencies, we intend for that guidance to also indicate what types of additional data or studies could have the greatest impact on improving the quality of the

assessment. In this way, the quality of criteria could continue to improve as data become available.

1.2 History of Aquatic Life Criteria

Prior to 1980, EPA derived aquatic life criteria for toxic chemicals using an "ad hoc" approach (i.e., formal procedures for their derivation were not codified). Criteria were usually established by citing the data deemed most relevant by those selected to derive the criterion for a given pollutant. This approach allowed for substantial inconsistencies in how toxicity data were used and the resulting level of protection provided, particularly since no minimum data requirements were established.

In 1980, EPA established for the first time written guidelines for deriving aquatic life criteria. These guidelines were last updated in 1985 (Stephan et al., 1985). In order to place the proposed approach for deriving tissue-based criteria into perspective, pertinent features of the 1985 Guidelines are summarized below.

Selected Features of the 1985 Guidelines

(1) Criteria are represented by a two-number system (an acute criterion derived for short-term exposures and a chronic criterion derived for long-term exposures) and are expressed as water concentrations.

(2) Species sensitivity is characterized using a species sensitivity distribution (SSD) with interpolation or extrapolation to obtain a criterion concentration protective of 95% of tested taxa.

(3) Minimum database requirements must be met in order to derive criteria (8 genera for acute criteria and 3 genera for chronic criteria from diverse taxonomic groups).

(4) Toxicity test data are based on water only exposures with only negligible exposure to chemicals from food.

(5) Acute criteria are based on 48-hr – 96 hr acute toxicity tests involving severe endpoints (e.g., survival, immobilization). Chronic criteria are based on longer term toxicity tests of early life stages, a partial life cycle, or a full life cycle involving endpoints such as survival, growth, reproduction and development. Data not conforming to these exposure durations are generally not used.

(6) Due to the limited amount of chronic toxicity data, derivation of chronic criteria often involves the use of acute-chronic ratios (ACRs) for extrapolating from acute to chronic effect concentrations.

(7) The 1985 Guidelines contain a procedure to derive a "final residue value" that attempts to address exposure to bioaccumulative chemicals. However, the science concerning bioaccumulation and subsequent EPA guidance for addressing bioaccumulation have evolved substantially over the last two decades such that this procedure is considered obsolete.

(8) The 1985 Guidelines also recommend an "averaging period" (1 hour for acute, 4 days for chronic) that is designed to address fluctuating exposures. The Guidelines also recommend an "allowable frequency" for exceeding the criterion (once in three years on average) which is intended to address the time needed for aquatic ecosystem recovery between criteria violations.

1.3 Limitations of 1985 Guidelines for Bioaccumulative Chemicals

In 1990, EPA convened a workgroup of scientists with the charge of revising the 1985 Guidelines to reflect the latest available science. Among other findings, ***the workgroup concluded that a separate set of procedures were needed for deriving aquatic life criteria for bioaccumulative chemicals.*** This conclusion grew out of recognition that the 1985 Guidelines contain a number of fundamental limitations with respect to deriving criteria for bioaccumulative chemicals. Specifically, the 1985 Guidelines:

(1) Lack a prescriptive procedure for addressing risks to aquatic life that result from exposure to chemicals from the diet (food web).

(2) Rely heavily on toxicity test data that often do not account for the slow accumulation kinetics of many bioaccumulative chemicals and consequently, may underestimate effects associated with long-term (steady state) accumulation.

(3) Lack a scientifically rigorous procedure for addressing chemical risks to aquatic-dependent wildlife (e.g., piscivorous birds and mammals).

1.4 Revision Efforts of the 1990s

Much of the effort of the Guidelines revision workgroup in the 1990s focused on developing a new framework for deriving aquatic life criteria for so-called "non-bioaccumulative" chemicals (i.e., chemicals where exposure via the diet is not a primary concern). As discussed in the companion SAB Consultation Document titled: *"Water-based Criteria,"* competing priorities impeded EPA's progress on the revising the 1985 Guidelines in the 1990s.

One of these competing priorities was the Great Lakes Water Quality Initiative (GLWQI) rulemaking, whereby EPA developed new chemical criteria for aquatic life, wildlife, and human health for the Great Lakes system (USEPA 1995a). Of particular relevance here is the GLWQI criteria focused on bioaccumulative chemicals and contained new procedures for deriving wildlife and human health criteria that accounted for chemical bioaccumulation in aquatic food webs[1]. These new procedures consisted of the use of bioaccumulation factors, biota-sediment accumulation factors and food web bioaccumulation models to estimate chemical accumulation in the aquatic diet of wildlife and humans residing in the Great Lakes system (USEPA, 1995b). These bioaccumulation methods were subsequently modified and extended to a national level with EPA's publication of its *Methodology for Deriving Water Quality Criteria for the Protection of Human Health* (USEPA, 2000; 2003). With appropriate modifications, the Subcommittee believes the bioaccumulation methods published in EPA's human health criteria methodology are applicable to aquatic life and aquatic-dependent wildlife receptors. To date, however, EPA has no national criteria methodology that specifically addresses risks from bioaccumulative chemicals to aquatic life and aquatic-dependent wildlife.

[1] The GLWQI criteria for aquatic life did not address food web bioaccumulation.

2 General Overview of Tissue-based Approach

2.1 What Are "Tissue-based Criteria?"

The Subcommittee is proposing to use a tissue-based approach for deriving criteria that protect aquatic life and aquatic-dependent wildlife from harmful exposure to bioaccumulative chemicals. We use the term "tissue-based criteria" to represent criteria that are derived from toxicological data expressed as concentrations in target organisms (e.g., commonly referred to in the literature as critical body residues, lethal body burdens, tissue residue-response relationships) as opposed to concentrations in ambient media (water, sediment). For aquatic-dependent wildlife, we consider tissue-based criteria to also include criteria that are based on toxicological data expressed as concentrations in aquatic organisms that compose their diet (e.g., mg chemical/kg food). The use of diet-based toxicological data will likely be reserved for aquatic-dependent wildlife because: (1) such data are more plentiful than toxicity data expressed as concentrations in wildlife tissues, and (2) exposure of wildlife to bioaccumulative chemicals from water ingestion is generally considered negligible relative to the diet, unlike aquatic organisms where exposure to chemicals via both food and water can be important. Thus, we use the term "tissue-based criteria" broadly to represent criteria derived both from toxicity data expressed as concentrations in tissues of the target organisms or their diet.

> We use the term "tissue-based criteria" broadly to represent criteria derived both from toxicity data expressed as concentrations in tissues of the target organisms or their diet (for wildlife).

2.2 Why Use a Tissue-based Approach for Bioaccumulative Chemicals?

The primary motive behind our pursuit of a tissue-based approach for bioaccumulative chemicals is the desire to account for multiple routes of exposure (e.g., diet, sediment, water) in the derivation of criteria. Chemical accumulation in the aquatic food web and subsequent dietary exposure is a dominant concern for bioaccumulative chemicals. Toxicological data based on internal (tissue) concentrations are attractive because they incorporate chemical uptake from different routes of exposure. Another motivating factor is that appropriate expressions of toxicity on a tissue concentration basis inherently account for toxicokinetic differences that exist among species. Conceptually, this should act to reduce variability in toxicity measurements between species caused by differing rates of uptake, distribution, metabolism, and elimination that would otherwise be reflected in media-based expressions of toxicity. Tissue-based expressions of toxicity also account for factors affecting the bioavailability of chemical concentrations in exposure media which can be a major source of variability in water-based toxicity test data. Finally, tissue-based expressions of toxicity appear to be promising for addressing exposure to chemical mixtures, particularly for those with common mode(s) of action.

The Subcommittee notes that the concept of expressing toxicological data for aquatic organisms on the basis of tissue or whole body concentrations is not new (e.g., Könemann, 1981) Veith et al. 1983; McCarty, 1986; Cook et al., 1989; 1993; McCarty and Mackay, 1993) and a substantial body of literature has evolved around this approach. For organic chemicals exhibiting a narcotic mode of action, the lethal tissue residue or body burden concept has its foundations in the early developments of quantitative structure activity relationships (QSARs)

involving octanol-water partition coefficients (K_{ow}), bioconcentration and acute toxicity (Veith et al., 1979; Veith et al. 1983; McCarty 1986). More recently, the lethal body-burden concept has been advanced as a method for deriving criteria for narcotic chemicals, including PAH mixtures (Di Toro, et al., 2000; Di Toro and McGrath 2000), although the toxicological basis for these criteria is driven mostly by measurements of acute lethality. EPA's development of the Biotic Ligand Model for cationic metals and application to deriving criteria for copper is also based implicitly on a lethal tissue residue approach (e.g., accumulation on the gill for fish; Di Toro et al., 2001; Paquin et al., 2002).

The use of tissue concentrations for expressing toxicological effects has been evaluated for compounds with reactive and specific modes of action (Verhaar et al., 1999; Legierse et al., 1999) and has been the subject of several critical reviews (Barron et al., 2002; Escher and Hermens, 2002; Beyer et al., 1996). For specific modes of action that involve irreversible (or less than reversible) binding of the toxicant with target sites, a critical target occupation model has been proposed for describing the time dependent toxicity (Legierse et al., 1999). This model does not assume a constant internal effect concentration with time, as often assumed with baseline toxicity. Furthermore, databases containing tissue-based toxicity data have been developed (USACE, 2004; Jarvenin and Ankley, 1999). Recently, EPA published draft aquatic life criteria for selenium that use a tissue-based approach (USEPA, 2004).

2.3 Guiding Principles of Tissue-based Criteria

The following principles or attributes helped guide the Subcommittee's thinking on how construct a methodology for deriving tissue-based criteria for bioaccumulative chemicals.

1. **Scientific Defensibility.** The methodology produces criteria that use the best available science and are scientifically defensible.
2. **Flexibility**. The methodology is flexible enough to accommodate the heterogeneous nature of available data including "data poor" and "data rich" situations.
3. **Transparency.** The methodology is transparent in how criteria are derived and how they can be set to satisfy different risk management objectives.
4. **Consistency.** The methodology is sufficiently prescribed such that its repeated application to the same dataset by appropriate users should result in the same (or similar) criteria values.
5. **Uncertainty.** The methodology does not discourage the generation of new data or methods for reducing uncertainty in the criterion.
6. **Site-Specificity.** The methodology is readily adaptable to enable derivation of criteria that reflect site- or region-specific attributes.
7. **Level of Effort.** The data requirements of the methodology are not be so onerous such that essentially no tissue-based criteria could be derived in the near future (i.e., the next 5 years) without the generation of a substantial amount of new data.
8. **Implementation.** The methodology facilitates the translation of tissue criteria to corresponding concentrations in various environmental compartments (e.g., water, food web) to address implementation and monitoring needs.

2.4 What is the General Process for Deriving Tissue-based Criteria?

This framework focuses primarily on a national-level process for deriving tissue-based criteria. The derivation of a national-level criterion will provide an analysis of all available toxicity data and a description and background on the parameter estimates used for representative species. National-level criteria may be adopted by State, Tribal, or local agencies or may be modified at state or local scales if sufficient additional information is available to improve the characterization of risk while maintaining the intended level of protection for aquatic life and wildlife. The framework will be expanded in the future to provide guidance on when and how site-specific criteria could be derived.

The current view of the Subcommittee is that guidelines for deriving tissue-based criteria for bioaccumulative chemicals would consist of two primary components:

1) Procedures for deriving a national tissue criterion (or criteria)

2) Procedures for translating a national tissue criterion (criteria) into corresponding concentrations in media and the aquatic food web.

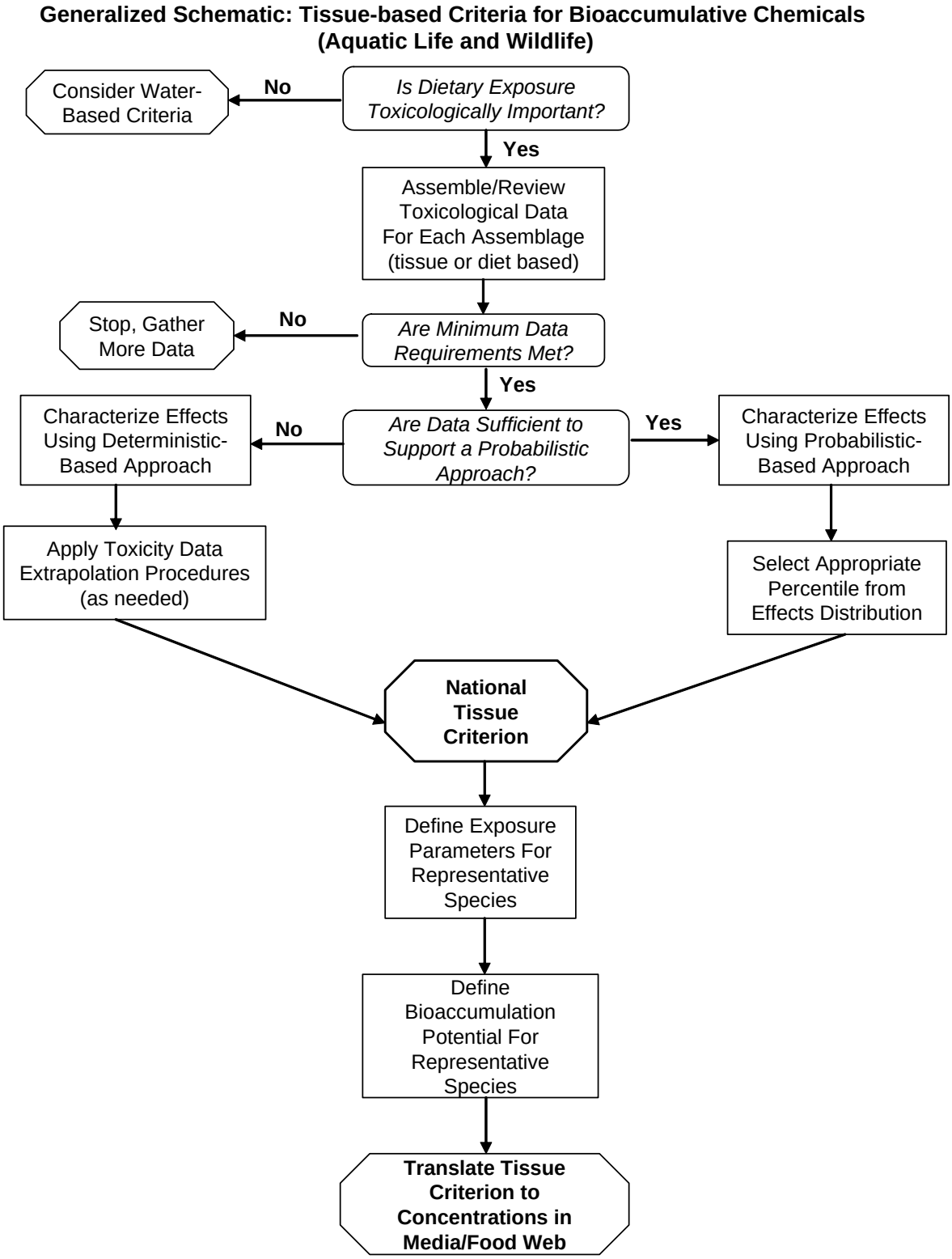
The second component (translating tissue criteria to media and food web concentrations) addresses both scientific and regulatory needs concerning the relationship between chemical loadings and accumulated chemical residues in tissue (i.e., bioaccumulation). Below, we provide an overview of these two components as they pertain to both aquatic life and aquatic-dependent wildlife with additional details following in Sections 3 and 4, respectively.

2.4.1 Derivation of a National Tissue Criterion

Figure 1 illustrates some of the primary decisions points and steps the Subcommittee is considering for deriving tissue-based criteria for aquatic life and aquatic-dependent wildlife with respect to bioaccumulative chemicals. For simplicity in presentation, we have chosen to represent only some of the decisions steps in the derivation process or have combined several steps into a single box.

As discussed in the Overview SAB Consultation Document, the derivation of a "water-based" or "tissue-based" criterion would begin with a "problem formulation" step, whereby critical assessment questions are formulated and addressed, a conceptual model developed, and an overall plan for analyzing the data is produced. Details of the problem formulation step are described in the Overview document and in Sections 3 and 4 of this document for aquatic life and aquatic-dependent wildlife, respectively. Therefore, it is assumed in Figure 1 that a complete problem formulation phase would be conducted, of which only certain steps are captured in this schematic. The primary steps of the schematic are as follows:

Figure 1.



1) **Determine Need for Tissue-based Criteria.** One of the first steps will be to determine the relative utility of a tissue-based approach for the chemical of concern as compared to a water-based approach. In the context of bioaccumulative chemicals, we expect the primary determinant to be the relative importance of chemical exposure via the diet to overall risks experienced by aquatic life and aquatic-dependent wildlife. Generally, the greater the importance of diet in governing chemical exposure and effects, the less likely that a water-based approach would be suitable and more likely a tissue-based approach would be used. Information on chemical properties (e.g, K_{ow}, persistence, etc.), trophic transfer, and toxicology of diet-borne chemical would be consulted as part of this decision step.

2) **Gather, Synthesize Toxicity Data.** Once an initial decision to pursue a tissue-based approach is made[2], all relevant data on the toxicity of the chemical would be assembled, with the primary focus on data that relate toxicological effects to chemical concentrations in tissue(s) and/or diet (in the case of wildlife). It is at this step in problem formulation where decisions are made about the form(s) of the chemical of concern, the most appropriate tissue(s) for expressing toxicological effects, the most appropriate toxicological endpoints to consider, relative sensitivity of taxa and life stages, in addition to proper screening of data for quality purposes.

3) **Determine Feasibility.** Once the appropriate toxicological data are reviewed, evaluated and synthesized, acceptable studies will be evaluated in the context of "minimum data requirements" which are established for each taxonomic assemblage to ensure that coverage of a diverse range of species is achieved. The assemblages being considered are aquatic plants, aquatic invertebrates, aquatic vertebrates (i.e., fish and some forms of amphibians), birds, mammals, and reptiles.

4) **Select General Derivation Approach.** If data meet the minimum data requirements, another key step within problem formulation will be to determine the approach to be used to characterize the effects on aquatic/wildlife species. The body of available toxicological data is evaluated to determine whether data are sufficient to support a probabilistic basis for setting the tissue criterion (e.g., species sensitivity distribution) or a deterministic basis (e.g., using data from a good quality study on an appropriately sensitive species). The selection of a probabilistic or deterministic approach is based on the quality and quantity of available toxicity data.

5) **Probabilistic Methods.** Probabilistic approaches have several advantages over deterministic approaches for deriving criteria, however they generally require that data be available from a relatively large number of species in order to reliably describe the overall distribution in species sensitivity. The decision on how much data are adequate to conduct a probabilistically-based approach has not been made by the Subcommittee but is one of many issues it intends to address in the future. In the case of species sensitivity distributions, one advantage is that the tissue criterion could be set using an appropriate

[2] Note: the tissue or water-based approaches are not mutually exclusive, and both may be initially applied for some chemicals where the relative value of one approach over the other is ambiguous.

percentile from the species sensitivity distribution. This is analogous to the current approach used in the 1985 Guidelines. An example of how a species sensitivity distribution can be applied to tissue-based toxicity data is illustrated by Steevens et al. (2005) for 2,3,7,8-TCDD.

6) **Deterministic Methods.** Deterministic approaches (e.g., characterizing effects based on data for an appropriately selected species or set of species) can be used with substantially less data but often are accompanied by the use of toxicity extrapolation procedures (e.g., uncertainty factors) which introduce uncertainty in the analysis.

7) **National Tissue Criterion.** Whichever approach is used to characterize toxicological effects, the goal of the methodology is to derive a tissue criterion (or criteria) that represents a concentration in tissue of aquatic life and/or aquatic-dependent wildlife that is deemed appropriately protective of the respective assemblages of species. In the case of aquatic-dependent wildlife, this tissue concentration may be expressed as concentrations in the aquatic diet.

2.4.2 Translation of Tissue Criterion to Concentrations in Media, Food Web

The Subcommittee anticipates the need to develop guidelines for translating tissue-based aquatic life and wildlife criteria into corresponding concentrations in environmental media (e.g., water) and/or other components of the aquatic food web for the following reasons:

- **Implementation.** Monitoring and enforcing pollutant discharge limits on the basis of measured chemical concentrations in tissues of organisms may not be practical or desirable in all situations (e.g., aquatic-dependent wildlife).

- **Intrinsic Toxicity vs. Risk.** The distribution of species sensitivity on the basis of tissue concentration-effect values (e.g., mg/kg-tissue) does not necessarily equate to the distribution of "risks" that would be experienced by those species from a given chemical concentration in water. While tissue-based toxicity data reflect the "intrinsic toxicity" of a chemical because bioavailability and toxicokinetic factors are addressed, such data do not reflect species-specific differences in exposure potential. For bioaccumulative chemicals, exposure potential can vary substantially among species due to differences in trophic position, habitat zone, and consumption rates. Therefore, the most sensitive species on a tissue concentration basis may not be the species "most at risk" on a water concentration (and chemical loading) basis.

The Subcommittee recognizes that translating tissue-based criteria for bioaccumulative chemicals into corresponding media concentrations involves a number of processes and parameters (e.g., bioaccumulation, food consumption patterns and rates, etc.) that can vary substantially across sites. Therefore, the current thinking is that procedures for translating tissue criteria into media concentrations would be developed in order to facilitate the use of appropriate site-specific data when available. In situations where such site-specific data are not available,

the Subcommittee anticipates that appropriate "nationally representative" parameter values could be used.

Continuing with Figure 1, the following general steps would be followed for translating tissue criteria into media concentrations.

1) **Define Exposure Potential of Representative Species.** It appears likely that the exact identity of species corresponding to the national tissue criterion (summarized above) will not be known. This situation is likely to occur because both deterministic and probabilistic approaches for characterizing effects will probably involve some type of extrapolation or interpolation of toxicity values among species (e.g., selecting a percentile from an SSD, applying uncertainty factors) in order to determine a tissue criterion that is protective of the overall assemblage. For example, the identity of a hypothetical species corresponding to the 5^{th} percentile from a SSD would likely be unknown, as would the components of its diet. Because the translation of tissue concentrations to media concentration requires knowledge of dietary composition, growth rates, feeding rates etc., we are proposing that a set of "representative species" be used to define exposure potential and the translation to media concentrations. Such species would be representative of the range of exposure potential likely to be encountered in the site(s) of concern, including "high end" exposure scenarios. Ideally, the representative species and associated exposure parameters (diet, body weight, food consumption rates, etc.) would be defined on a site or regional basis. In situations where this is not possible, the Subcommittee envisions that a "default" set of nationally representative species and parameter values would be developed. For aquatic life, these species would reflect a range of feeding guilds (e.g., carnivory, piscivory, omnivory, herbivory), habitat preferences (e.g., benthic, pelagic), and taxonomic groups within each assemblage. Similarly for wildlife, a set of representative species would reflect a range of feeding guilds, taxonomic groups, and habitat types across the United States.

2) **Define Bioaccumulation for Representative Species.** Once the representative species have been defined and exposure parameters characterized (either on a site or national basis), the next step is to define the bioaccumulation potential for the chemical in the context of each representative species. Since chemical bioaccumulation in aquatic organisms can vary on across sites, bioaccumulation would ideally be characterized using site-specific information. For nonionic organic chemicals, some key factors include disequilibrium between chemical concentrations in sediment and water, lipid content, dissolved and particulate organic carbon, food web structure, trophic position, metabolism, and hydrophobicity. The Subcommittee envisions using a combination of empirical (e.g., field-derived bioaccumulation factors, biota-sediment accumulation factors) and mechanistic models (e.g., food web bioaccumulation models) for assessing a chemical's bioaccumulation potential. For situations where a site-specific assessment of bioaccumulation potential is not possible, the Subcommittee is considering the need to derive a set of nationally representative bioaccumulation factors (BAFs) that could be used to characterize bioaccumulation potential. This appears to be most applicable to organic chemicals where factors such as lipid fraction and dissolved and particulate organic carbon that can be readily adjusted to reflect local or regional conditions. This

approach is consistent with EPA's bioaccumulation assessment guidance developed for deriving human health water quality criteria (USEPA, 2000, 2003).

3) **Translation to Media Concentrations.** For aquatic life, translation of the tissue criterion to corresponding water concentrations would be accomplished by dividing the tissue criterion by the appropriate BAF derived for each representative species. An analogous approach could be constructed for translating to sediment concentrations. For wildlife criteria derived from dietary toxicity data, BAFs would be applied and appropriately weighted for each component of the aquatic diet of the representative wildlife species.

2.5 Challenges to Deriving Tissue-based Criteria for Bioaccumulative Chemicals

Basic toxicological principles suggest that measurements of exposure closer the site(s) of toxic action (e.g., tissue or body residues) is preferred over measurements in external media (water). In practice, however, a number of factors can act to mitigate the conceptual advantages of tissue-based criteria over water-based criteria. Some of these include:

1. The scope and quantity of applicable toxicological measurements based on tissue concentrations appears far more limited compared to water-based measurements. Given that aquatic life criteria are intended to protect entire aquatic communities from harmful exposures, a reduction in the number of species from which to estimate such criteria generally translates into greater uncertainty associated with the criterion.

2. Related to #1 above, the applicability of existing tissue-based toxicological measurements for criteria derivation appears to vary substantially. A sizable portion of the tissue-based toxicity data compiled to date reflects measurements of chemical concentrations in multiple types of tissues (even within the same study) in combination with a given toxicological response. Notably, the mere measurement of a chemical concentration in tissue(s) in tandem with a toxicological effect does not solely constitute a valid toxicological linkage between a given tissue concentration and an associated effect. Of critical importance for making toxicological inferences is establishing a valid tissue concentration-response relationship for appropriate tissues in conjunction with an understanding of the mode(s) and site(s) of action.

3. Ambiguity in tissue concentration-response relationships can also result from incomplete knowledge of the bioavailable form(s) of chemicals in tissue (particularly problematic with metals; Rainbow 2002), the effect of exposure route on the potency of a given tissue concentration, and even duration of exposure (e.g., Landrum et al., 2004).

3 Proposed Process for Deriving Aquatic Life Criteria for "Bioaccumulative" Chemicals

3.1 Importance of Problem Formulation

As described in the SAB Consultation "Overview Document," problem formulation is the initial step in a risk assessment where information about the chemical stressor, its exposure potential, and its effect on the ecological receptors of concern is evaluated for defining the scope of the assessment and for ensuring that the risk management goals are met. We believe most, if not all, elements of problem formulation are relevant to the derivation of aquatic life criteria. It is in the problem formulation step where the decision to apply a tissue-based approach is made. Assessment questions are formulated and addressed, important data gaps are identified, and a conceptual model is developed. Importantly, a plan is devised for analyzing the data and formulating the criterion that makes best use of the available information. This analysis plan is particularly relevant to tissue-based aquatic life criteria since a flexible approach is being proposed for deriving criteria depending on the availability of data and assessment needs (Figure 1). The summary below presents the current thinking of the Subcommittee regarding several important issues related to deriving tissue-based aquatic life criteria for bioaccumulative chemicals.

3.1.1 Deciding Between a Tissue or Water-based Approach

The Tissue-based Criteria Subcommittee is focusing on developing criteria for chemicals for which water concentration is not a reasonable surrogate for target tissue toxicant concentration expected under natural exposure conditions. In other words, the Subcommittee is focused on chemicals for which water concentration does not adequately capture exposure and subsequent toxicological effects expected in the natural environment. Such chemicals generally bioaccumulate extensively in aquatic food webs such that trophic transfer and subsequent dietary exposure become toxicologically important. Organic chemicals in this category generally have high hydrophobicity (e.g., log K_{ow} > 5), long environmental persistence, and are poorly metabolized by biota. A few obvious examples include polychlorinated dioxins, furans, and biphenyls, DDT & metabolites, and dieldrin. Selected organometallics and metalloids also fall into this category (e.g., methylmercury, selenium).

In many cases the decision to pursue a tissue-based approach will be obvious from the onset. All relevant information on the toxicological importance of dietary exposure will be considered. For some chemicals, however, the Subcommittee expects this decision to be ambiguous (e.g., perhaps for some organic chemicals with log K_{ow} values in the 4-5 range). In such cases, both a water and tissue-based approach may be pursued with a final decision being based on the relative uncertainty among the two approaches. Furthermore, the Subcommittee notes that the relative importance of dietary exposure can vary widely across species for a given chemical. Some groups of organisms with high food intake rates and high chemical assimilation efficiencies (e.g., high volume filter feeders) may be especially prone to chemical exposure via the diet. If such organisms are among the most toxicologically sensitive to the chemical in question, then they may be particularly relevant in the decision to use a tissue-based approach. In addition to direct evidence of the relative toxicological importance of dietary exposure,

indirect evidence via bioaccumulation modeling involving multiple exposure pathways may be considered. The Subcommittee emphasizes that the toxicological importance of dietary exposure is key (not just the existence of dietary exposure), as some organisms may be highly exposed via the diet but have evolved storage and detoxification mechanisms that can render the toxicological importance of accumulated chemical concentrations in tissue as being minimal or ambiguous (e.g., selected marine invertebrates; Rainbow, 2002).

3.1.2 Addressing Key Assessment Questions

The problem formulation step is where assessment questions are formulated and addressed. Examples of assessment questions are provided in the SAB Consultation "Overview Document." Highlighted below are several assessment questions and issues the Subcommittee has discussed to date in the context of deriving tissue-based criteria for bioaccumulative chemicals.

Understanding Mode(s) of Action. Understanding the mode(s) of action is important for a number of reasons. First, information on mode of action can aid in distinguishing among taxonomic groups in terms of their expected sensitivity, particularly when combined with information on key physiological attributes (e.g., presences/activity of AhR receptors for exposure to dioxin-like compounds). Knowledge on mode of action can also be important for interpreting tissue concentration-based toxicity data. Specifically, the relative reversibility/irreversibility of the mode of action may aid in understanding the importance of exposure duration in affecting the potency of chemical concentrations in tissue. For example, there is evidence in the literature that some specific modes of action may involve irreversible (or less than reversible) binding to toxicological receptors such that the potency of a given tissue concentration increases with increasing exposure time (Lee et al., 2002a; 2002b; Landrum et al., 2004; 2005). This in turn may affect how one chooses to aggregate tissue-based toxicity data or conduct toxicity data extrapolations. Mode of action information is also important in the decision to derive criteria on the basis of chemical mixtures. The Subcommittee notes that there may be ambiguity in identifying the critical mode(s) of action or limitations to making inferences based on mode of action data because: (1) most mode of action data for aquatic organisms have been gathered from acute toxicity tests involving fish, (2) mode of action might vary across species, life stages and with the magnitude and duration of exposure, (3) multiple or unknown modes of action may be involved with the expression of toxicological effects.

Understanding Potency of Tissue Concentrations Derived from Different Exposure Routes. A critical issue for interpreting tissue concentration-based toxicity data is how one addresses the potency of chemical concentrations in tissues that are derived from different exposure routes (e.g., water vs. food). One of the most attractive features of a tissue-based approach is the notion that toxicity from different exposure routes can be integrated by a chemical concentration measured in an appropriate tissue. If this were not the case, then the utility of tissue concentration-based toxicity data would be significantly compromised due to the highly heterogeneous nature of toxicity test designs (e.g., exposures from water, sediment, food, injection). For organic chemicals that obviously fall into the

"bioaccumulative" category, the Subcommittee is not aware of evidence that tissue concentration-based toxicity values routinely vary by exposure route. For metals, there is evidence that the route of exposure can affect the potency of a given concentration in tissue. However, most metals would not be considered in the context of criteria for bioaccumulative chemicals. The Subcommittee invites SAB comment on the importance of exposure route in affecting toxicity expressed as concentrations in tissue, particularly with regard to organic chemicals.

Understanding the Importance of Temporal Variability in Exposure Concentrations. As described in the SAB consultation document on Water-based Criteria, modeling toxicity as a function of short-term (daily) fluctuations in water concentrations is a fundamental component of the proposed water-based criteria methodology. For bioaccumulative chemicals (e.g., persistent organic chemicals w/ log K_{ow} >5), the current thinking of the Subcommittee is that such short-term fluctuations will generally be much less important in affecting chemical uptake and tissue concentration-based toxicity. The basis for this thinking originates in the notion that for most aquatic species of concern (e.g., especially larger bodied animals at higher trophic levels such as piscivorous fish), accumulation kinetics of "bioaccumulative" chemicals is sufficiently slow such that risks from short-term (acute) exposures are generally not nearly as important relative to risks from long-term exposures. An illustration of this phenomenon is shown in Figure 2 using tissue concentrations predicted by the Gobas (1993) food web bioaccumulation model. It can be seen in Figure 2 that concentrations of highly hydrophobic chemicals (e.g., log K_{ow} > 5) in piscivorous fish are dampened temporally compared to concentrations in water. As a result, the Subcommittee expects that tissue-based aquatic life criteria for bioaccumulative chemicals will be concerned with chronic exposures and conditions approximating steady state. However, the Subcommittee recognizes that exceptions to this generalization might occur, possibly for small-bodied organisms lower in the food web where accumulation kinetics might be relatively rapid (e.g., zooplankton). If such organisms are among the most sensitive species to the chemical in question, then steps to address risks associated with short-term exposures will need to be taken. This might involve using dynamic bioaccumulation modeling for translating critical tissue residues back to media concentrations and/or the use of shorter-term averaging periods per the 1985 Guidelines methodology.

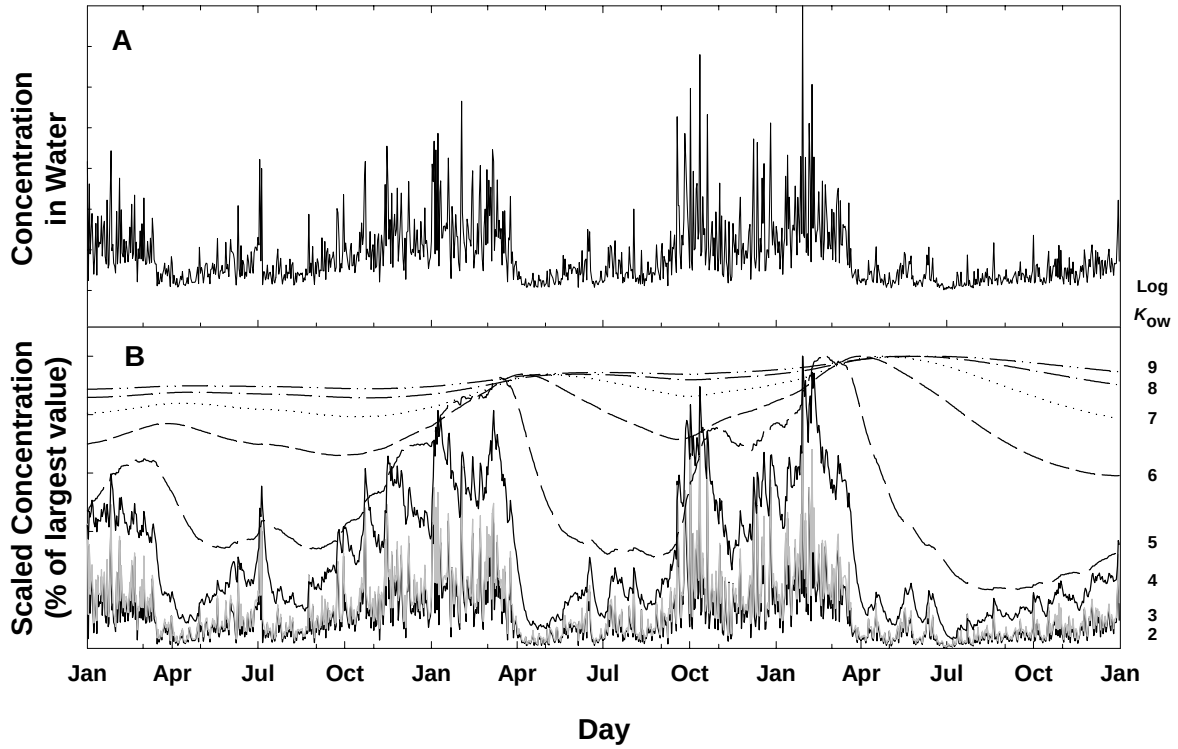


Figure 2 (A). Daily concentrations of a hypothetical nonionic organic chemical over time in the water column, predicted using a simple dilution model and daily flow data for the Mississippi River at St. Paul, Minnesota. **(B)** Daily chemical concentrations in piscivorous fish found using the kinetic food web models of Gobas (1993) with the daily chemical concentrations in the water column for nonionic organic chemicals with log *n*-octanol-water partition coefficients (log K_{ow}s) of 2, 3, ... and 9. The daily chemical concentrations in piscivorous fish have been scaled to the largest value for each log K_{ow}. As hydrophobicity increases, temporal variability in chemical concentrations in piscivorous fish decreases dramatically. From Burkhard (2003)

3.1.3 Screening and Evaluation of Toxicity Data

An important step in the derivation of tissue-based aquatic life criteria for bioaccumulative chemicals involves the screening and evaluation of toxicity data. All toxicity data will be first reviewed for acceptability based on quality prior to their use in deriving a criterion. Most of the evaluation criteria used to determine acceptable toxicity test quality will be the same or similar to those used in the derivation of water-based aquatic life criteria. A few examples include:

- Sufficient written documentation must be available from which to judge the quality of the methods, measurements, and statistical analyses conducted. Peer reviewed publications are preferred.
- Laboratory studies must contain a control treatment with an acceptable response rate for control organisms (often specific to the test design).
- Test organism handling, holding, acclimation, and loading rates should conform to standard practices appropriate to the test design (e.g., ASTM or other similar peer reviewed guidelines).

- Water quality parameters (e.g., pH, temperature, D.O., etc.) and their rates of change must be within accepted ranges for test design and/or the environmental tolerances of the test organisms.

In addition to data quality considerations, the Subcommittee recognizes that a number of other attributes of tissue concentration-based toxicity data will likely need to be evaluated for determining their acceptability and/or utility for deriving aquatic life criteria. Some of these attributes are listed below.

- **Study Duration.** The primary focus for characterizing effects of "bioaccumulative" chemicals will be those toxicity data that are indicative of effects resulting from long-term (chronic) exposure. This focus is based in part on the notion that risks will likely be driven from chronic exposure not from acute exposures. Furthermore, mode of action may vary as a function of magnitude and duration of exposure. We note, however, that certain tissue concentration-based toxicity data derived from short-term exposures are appropriate. For example, short-term exposure to critical life stages (e.g., egg, embryo) form the basis of estimating risks of 2,3,7,8-TCDD to aquatic life (Cook et al., 1993). The Subcommittee recognizes that toxicity data expressed on the basis of tissue concentrations may have the capacity to integrate differences in the magnitude and duration of chemical exposure, particularly for non-specific, reversible modes of action (e.g., narcosis). However, for some specific modes of action that involve irreversible or partially irreversible binding to receptors involved in expressing toxicity, both theoretical and empirical evidence suggest that exposure duration can contribute to variance in tissue-based toxicological effect levels. For example, exposure duration has been suggested as a source of variability observed in the potency of PAH concentrations in tissue (Lee et al., 2002a; 2002b). Duration of exposure may be particularly important for chemicals which are slow to reach steady state with respect to their distribution in organism tissues. Therefore, the effect of exposure duration on the potency of a particular tissue concentration is an attribute that would receive specific evaluation for characterizing effects using tissue-based toxicity data.

- **Toxicological Endpoints.** Similar to the 1985 Guidelines, studies will be screened for those that measure effects based on toxicological endpoints that are thought to have most relevance to potential impacts on populations (e.g., survival, growth, reproduction, development). Other endpoints would be used provided that a sufficient ecological and toxicological linkage can be made to impacts on populations.

- **Strength of Tissue Concentration-Response Relationship.** Analogous to the evaluation of water-based toxicity data, the strength of the tissue concentration-response relationship is an important criterion for evaluating the acceptability of tissue-based expressions of toxicity. This point cannot be overemphasized for evaluating tissue-based toxicity data, since often times chemical concentrations are analyzed and reported for multiple tissues along with a common set of adverse effects. Chemical concentrations in some tissues may have little or no correlation with toxicological effects. However, this lack of correlation may be reflect confounding factors such as the use of too few

treatments or treatment levels that are beyond the range of the concentration-response curve.

In order to be used directly in the derivation of a tissue-based criterion, a "valid" tissue concentration-response relationship must be available from the toxicity test in question. Validity of the tissue concentration-response relationship will likely be judged quantitatively based on its statistical and toxicological significance and qualitatively based on consistency of increasing response with increasing tissue concentrations. Tissue concentration-response relationships that can be expressed quantitatively (e.g., via regression equations) are generally preferred over those that can only be expressed qualitatively (i.e., increasing response with increasing tissue concentration). Tissue concentration-effects data for which a concentration-response relationship is not observed may have some utility for characterizing effects (e.g., in the case of unbounded NOAELs, indicating levels where effects have not occurred).

- **Tissue Type.** The choice of tissue(s) used to relate chemical concentrations to toxicological effects is an important attribute to consider when developing tissue-based criteria. Other factors aside, preference will generally be given to tissues that either represent or are closely linked to the site(s) of toxic action. However, the choice of tissue(s) upon which to base the effects characterization will have to balance proximity to the site(s) of action with the availability of data for that tissue and the ability to extrapolate chemical concentrations between tissues. For example, cursory examination of the tissue concentration-based toxicity data indicates that the preponderance of data exists in the form of whole body concentrations (Appendix A). Thus, it appears that whole body concentrations will need to be used as surrogates for concentrations at the site(s) of toxic action and/or integrated with models for estimating concentrations in specific tissues.

3.1.4 Minimum Data Requirements and Assemblages

The current view of the Subcommittee is that tissue concentration-based toxicity data will need to be evaluated against a set of "minimum data requirements" (MDRs) before a criterion could be derived. This concept is consistent with the 1985 Guidelines and the new proposal for deriving water-based criteria. Minimum data requirements are a defined set of taxonomic or ecologically-based species groups from which acceptable toxicological data must be available in order to derive a criterion. The Subcommittee believes some set of MDRs are needed in order to preserve a minimum level of reliability in tissue-based aquatic life criteria.

The Subcommittee has discussed the issue of MDRs but has not reached final consensus on an exact set of MDRs to propose. However, current thinking is that MDRs would be defined separately for three assemblages of aquatic organisms:

(1) vertebrates,
(2) invertebrates
(3) plants

These three assemblages are consistent with the current proposal for deriving water-based aquatic life criteria (see SAB Consultation Document on Water-based Criteria). Defining MDRs separately for different assemblages might allow for some flexibility when deriving criteria. For example, criteria might be derived only for those assemblages for which sufficient data are available. Although the exact composition of MDRs has not been specified, current thinking is that they would consider taxonomic diversity in addition to factors related to a species' "ecological niche" as defined by trophic status/feeding guild, habitat preference, life history, etc. Thus, MDRs could be defined as requiring data for a top predatory (piscivorous) fish, a benthic feeding carnivorous fish, an herbivorous fish, etc. Life history attributes such as generation time (an important influence on population recovery time) may also be considered so that both short-lived and long-lived species would be represented. The Subcommittee also recognizes practical constraints to defining MDRs. If MDRs are too onerous, few if any criteria could be derived in the near future. Based on a cursory review of the availability of tissue-based toxicity data for aquatic organisms, it appears that MDRs for deriving deterministically-based criteria would approximate 4-5 species per vertebrate and invertebrate assemblage. The Subcommittee has not discussed MDRs for plants.

3.1.5 Deciding Between a Deterministic or Probabilistic-based Effects Characterization

The MDRs discussed above would presumably apply to deriving criteria using a deterministic-based approach (i.e., that approach requiring the least amount of toxicity data). However, if sufficient data were available, probabilistic-based methods for characterizing effects would be considered (e.g., SSD, the TEA model described in the SAB Consultation Document on Water-based Criteria). In cases where the advantages of a probabilistic approach over a deterministic approach are not obvious, the Subcommittee can envision that criteria would be derived using both general approaches, with the approach that achieves the management goals with the least uncertainty becoming the preferred method. The Subcommittee has not discussed the quality or quantity of data required to apply a probabilistic-based approach such as a SSD but will address this issue in the future.

3.2 Characterization of Effects: Deterministic Criteria

The focus of the effects characterization is quantifying relationship between accumulated chemical concentrations in tissues and toxicological effects across multiple species in order to support the selection of a chemical concentration in tissue that would adequately protect a given assemblage. In the ideal situation, an abundance of toxicological data would be available from which quantitative relationships between tissue concentrations and adverse effects could be established. In the aggregate, the underlying toxicological database should ideally represent:

(1) a diverse array of aquatic species (e.g., multiple families of fish, invertebrates and perhaps plants) in order to capture variability in sensitivity among species within each assemblage,
(2) a diverse array of toxicological endpoints that can be closely linked to population-level effects (e.g., reproduction, mortality, growth, development to name a few),
(3) chemical measurements in tissues that represent, or are closely linked to, the site(s) of toxic action, and

(4) toxicity tests conducted under standardized protocols with regard to routes of exposure, duration, life stages tested, etc.

In reality, the situation appears to be far from this ideal. Cursory examination of two compilations of tissue-effects data (Jarvenin and Ankley, 1999; USACE, 2004) reveals that the majority of chemicals have relatively few species represented, a strong dominance of lethal endpoints over sublethal endpoints, a variety of routes of exposure, and most measurements in whole body vs. specific tissues (see Appendix A). Perhaps the greatest obstacle facing the successful derivation of tissue-based aquatic life criteria for bioaccumulative chemicals is the relative lack of appropriate, standardized, tissue-based toxicological data.

The following sections provide some insight into the thinking of the Subcommittee on how tissue concentration-based toxicity data would be synthesized for supporting a deterministically-based criterion.

3.2.1 Characterizing Effects on Organisms

The overall goal in this step is to define concentrations in tissue(s) below which unacceptable adverse effects on the test organisms are not likely to occur. Some of the Subcommittee's general preferences for synthesizing toxicity data among studies within a species are provided below.

- In general, determination of tissue-based effect concentrations using point estimation methods (e.g., ECxx based on regression analysis) is preferred over those determined by hypothesis testing (e.g., ANOVA-based NOAELs and LOAELs). All else being equal, point estimation methods enable interpolation between treatment levels to obtain a more precise estimate of the magnitude of effect compared to hypothesis testing methods.
- Studies with treatments (or observations in the case of field data) that bracket the onset of unacceptable adverse effects are preferred over those studies where either: (1) all treatments showed unacceptable adverse effects, or (2) no treatments showed unacceptable adverse effects.
- Defining what constitutes an unacceptable adverse effect (i.e., the magnitude of effect or EC_{xx}) will likely depend on the toxicological endpoint measured. Results of population modeling could conceivably help inform the selection of an appropriate EC_{xx} (see Section 3.2.3 and the population modeling discussion in the SAB Consultation Document on Water-based Criteria).
- Studies using a chronic exposure duration involving multiple life stages (or exposure to early or other critical life stages) are generally preferred over those of shorter exposure duration involving single life stages.
- If two or more acceptable tissue-based effect concentrations are available for a given species, life stage, and endpoint (e.g., mortality), the study that is considered to be of the highest quality and containing the least uncertainty in quantifying the threshold for unacceptable effects would be selected. Likely factors to consider in this evaluation include:
    a) environmental realism of exposure regime
    b) statistical power of the study

c) statistical uncertainty associated with the tissue-based effect concentrations
d) repeatability of the test results
e) accuracy and precision of the biological and chemical measurements
f) uncertainty associated with extrapolating results to the field.

- If no discernable difference exists between the quality and uncertainty associated with two or more studies involving the same species, life stage and endpoint, current thinking is that tissue-based effect concentrations would be averaged. This would help minimize the impact of inter-test variability on selecting a representative tissue-based effect concentration for a given species.

If two or more acceptable tissue-based effect concentrations are available for the same species and endpoint but for different life stages, preference would be given to the values from more sensitive life stage(s) for characterizing effects on that species (unless data are being used in population modeling where data for multiple life stages are preferred).

3.2.2 Toxicity Data Extrapolations

The Subcommittee expects that limitations in the scope and quantity of tissue-based toxicity data will require that various extrapolations be made in order to derive aquatic life criteria that can achieve an adequate level of protection. The Subcommittee has identified various types of toxicity data extrapolations that may be needed (below) but has had very little discussion to date on how to conduct such extrapolations. The Subcommittee invites SAB comment on the need and methods for conducting toxicity data extrapolations on a tissue concentration basis.

- **Extrapolating Across Magnitudes of Effect.** The Subcommittee envisions a potential need for extrapolating from higher magnitudes of effect to lower magnitudes of effect (e.g., LOAEL to NOAEL, EC_{50} to EC_{10}) in cases where tissue concentrations corresponding to lower magnitudes of effect are not quantified or are not reported for a given endpoint (e.g., mortality). Statistical modeling may be used in cases where the tissue concentration-response relationship has been adequately defined. In cases where the tissue concentration-response relationship has not been adequately defined, traditional approaches for human health and wildlife criteria have used uncertainty factors (UF). Methods for developing or selecting UFs have not been discussed by the Subcommittee.

- **Extrapolating Across Exposure Duration.** In cases where there is sufficient evidence to indicate the potency of a given chemical concentration in tissue is influenced by exposure duration, it is conceivable that some type of extrapolation may be needed to relate observed effects from shorter exposure durations to those expected from longer (chronic) exposure durations. There is some evidence of time-dependent toxicity of tissue concentrations in the literature for certain compounds (Lee et al., 2002a; 2002b; Landrum et al., 2004; 2005). However, these studies involve relatively short exposure durations (10 days or less), and their applicability to longer-term chronic and subchronic exposures (which is the general focus for bioaccumulative chemicals) is not clear. If chronic and subchronic tissue-based toxicity data are subject to time-dependency, the use of tissue concentration-based toxicokinetic modeling may be required.

- **Extrapolating Between Tissues.** The Subcommittee expects a need to extrapolate between tissues for expressing tissue concentrations associated with adverse effects (for example, from concentrations in whole body to concentrations in specific tissues). This may be required to place available tissue-based toxicity data on a common basis. For highly hydrophobic organic chemicals, current thinking is that information on lipid content of different tissues may be used for extrapolating tissue-based effect levels between tissues. In other situations, use of empirical relationships may be required for relating chemical concentrations between tissues.

- **Extrapolating Between Species.** Assuming that deterministic-based criteria could be derived with as few as 4-5 species within an assemblage, the current thinking of the Subcommittee is that some type of interspecies extrapolation of toxicity may be needed to account for untested species of an assemblage that may be substantially more sensitive than the most sensitive species tested. The assumption here is that we would likely be addressing specific modes of action where species sensitivity can differ substantially on a tissue concentration basis as opposed to nonspecific modes of action (narcosis) where effects may be more narrowly distributed on a tissue concentration basis. In the context of a cumulative frequency distribution, the most sensitive species among a dataset containing four species approximates the 25^{th} percentile, a level substantially larger than traditional aquatic life criteria which are set at the 5^{th} percentile. However, because aquatic life criteria that are derived using the 1985 Guidelines combine data from the aquatic invertebrate and vertebrate assemblages into a single SSD for as few as 8 species, this comparison is not entirely straightforward. The Subcommittee notes that methods for extrapolating toxicity between species have been derived from toxicity data expressed as concentrations in exposure media (e.g., ICE, Asfaw et al., 2003). However, we are not aware of methods for interspecies extrapolation of toxicity on a tissue-concentration basis. At this point in time, the Subcommittee has not discussed how such extrapolations would be conducted with tissue concentration-based toxicity data and solicits SAB comments on the issue.

3.2.3 Characterizing Effects on Populations

The Subcommittee on Tissue-based Criteria has not had detailed discussions on characterizing the effects of bioaccumulative chemicals at the population level for aquatic organisms, deferring instead to the expertise and work in this area being conducted by the Water-based Criteria Subcommittee. Conceptually, population models being considered for deriving water-based criteria should be applicable to tissue-based toxicity data available for bioaccumulative chemicals. For example, Munns et al. (1997) used a stage-specific, density independent model to estimate the effects of dioxin and PCB tissue concentrations on the intrinsic rate of population growth for the mumichog, *Fundulus heteroclitus.*

In practice, however, the feasibility and utility of population modeling appears ambiguous to this Subcommittee in the context of tissue-based criteria for bioaccumulative chemicals. Part of this ambiguity relates to the apparent limited availability of tissue concentration-based toxicity data for multiple life stages within a species. Lacking data to

characterize the differential sensitivity of different life stages would appear to significantly limit the ability to parameterize stage-specific population models. Furthermore, the availability of tissue-based toxicity data for reproductive endpoints appears to be extremely limited, based on a review of two databases containing tissue concentration (residue)-based toxicity information. Finally, the utility of population modeling in the context of constant (time invariant) exposure concentrations is also questionable to the Subcommittee. Part of the rationale for using population models for water-based criteria is to characterize effects resulting from fluctuating exposure concentrations and to integrate recovery time. If toxicity modeling for bioaccumulative chemicals is generally limited to constant concentrations in tissue (steady-state conditions) as discussed earlier in this proposal, the "value added" of population modeling appears, at least at this point, to be unclear. The Subcommittee plans to conduct additional analyses to clarify the role and utility of population modeling for setting aquatic life criteria for bioaccumulative chemicals.

3.2.4 Setting a Deterministically-based Tissue Criterion

The goal of a national tissue criterion for aquatic life would be to represent a concentration in tissue that at or below which the likelihood of unacceptable adverse effects on aquatic life would be appropriately low (i.e., as determined by risk management goals). Where multiple criteria are derived for different assemblages (e.g., invertebrates, vertebrates, plants), current thinking is that criteria for the most sensitive assemblage would apply due to the interdependence among assemblages in maintaining healthy ecosystems. Within an assemblage, current thinking is that the tissue criterion would be derived from data for a species that enables the protection goals to be met with the least uncertainty. Generally, this will be the most sensitive species. However, exceptions may exist in cases where uncertainty associated with basing the tissue criterion on the data for the most sensitive species is considered substantially higher than basing the criterion on data from a less sensitive species (e.g., the next most sensitive species). As discussed in Section 3.2.2, some type of toxicity data extrapolation may be needed to address concerns over the potential for greater sensitivity of untested species. The technical basis for conducting this extrapolation or evaluating uncertainty has not been discussed by the Subcommittee.

3.3 Characterizing Effects: Probabilistic-based Criteria

One option being explored by the Subcommittee for characterizing effects on a probabilistic basis involves the use of Species Sensitivity Distributions (SSD). Characterizing effects on the basis of SSDs forms the foundation of the 1985 Guidelines. One distinct advantage of an SSD approach over the deterministic approach described above is that the criterion can be selected to conform to a specified "risk level" or percentile (e.g., setting at a 5^{th} percentile to theoretically protect 95% of the tested species, per the 1985 Guidelines) via interpolation or extrapolation using statistical techniques. Aside from enabling consistency in the "risk level" selected across chemicals with heterogeneous datasets, the use of statistically-based interpolation or extrapolation techniques with the SSD approach enables one to mitigate

the influence of potential "outliers" on the derivation of the criterion, at least when compared to selecting the most sensitive species as the basis of the criterion.

A significant obstacle to the use of SSDs for tissue-based criteria is that they require data for a relatively large number of species in order to characterize species sensitivity with statistical rigor. To be statistically valid, SSDs should ideally be composed of data for the same or similar toxicological endpoints. If the underlying toxicity test data lack consistency in test design and endpoints measured, the SSD derived from such data would not only reflect true sensitivity differences but also differences related to test design. The Subcommittee has not discussed which specific SSD models it would recommend for use nor criteria for judging when to apply an SSD for deriving a tissue criterion.

The Subcommittee has also not discussed in detail the feasibility of applying the Toxic Effect Aggregation model (TEA) for characterizing effects, which is described in the SAB Consultation Document on Water-based Criteria but plans to do so in the future as details with the TEA model are resolved.

3.4 Setting a Probabilistically-based Criterion

Assuming that a valid SSD could be constructed using tissue-based toxicity data, then a criterion value could be selected to correspond to any desired level of 'risk' (i.e., any percentile of the SSD). For example, the 1985 Guidelines set criteria to correspond to the 5th percentile of the SSD. To date, we have not discussed a specific percentile or range of percentiles at which to set the national tissue criterion. However, in order to facilitate different risk management options to be considered, the Subcommittee is promoting transparency and flexibility in the selection of the percentile(s) for setting a national tissue criterion rather than setting it at a single percentile specified *a priori*.

3.5 Translating Tissue Criteria to Concentrations in Water, Food Web

As discussed in Section 2, the Tissue-based Criteria Subcommittee anticipates the need to develop guidelines for translating tissue-based aquatic life criteria into corresponding concentrations in environmental media (e.g., water) and relevant components of the aquatic food web. Translating tissue-based criteria into concentrations in ambient environmental media is often required for implementing criteria through regulatory programs. Translating tissue-based criteria into corresponding concentrations in components of the aquatic food web may also be required to facilitate monitoring of tissue concentrations (e.g., monitoring chemical concentrations in the diet of fish may be more practical than direct monitoring of fish tissue in some cases). This section presents the thinking of the Subcommittee regarding how national tissue-based criteria for bioaccumulative chemicals might be translated to other compartments of the aquatic ecosystem.

Ideally, this translation would be conducted using data specific to the site(s) of concern because many attributes can affect bioaccumulation of chemicals on a site-specific basis (e.g.,

food web structure, organic carbon concentration, chemical disequilibrium between sediments and water, etc.). Therefore, consistent with other EPA guidelines on estimating chemical bioaccumulation (USEPA 2000; 2003), use of site-specific data for translating tissue criteria to media concentrations would be strongly encouraged. However, past experience indicates that site-specific data (and/or the resources to obtain such data) may not be available in some circumstances. Therefore, the Subcommittee is considering the possibility that a default set of conditions may have to be defined for translating tissue criteria into media concentrations for use in circumstances where site-specific data are unavailable. This approach is consistent with past EPA guidance on bioaccumulation.

3.5.1 Use of Representative Species

Translating a tissue-based aquatic life criterion to media concentrations would initially appear to be a straight-forward task that would involve the use of bioaccumulation models. However, the following two issues arose during the Subcommittee's discussion of this translation step.

- **Ambiguity in Species Identity Associated with a Tissue Criterion.** If the tissue criterion were derived with the use of extrapolation or interpolation techniques (e.g., uncertainty factors for deterministic criteria; at a specified SSD percentile for probabilistic criteria), the identity of the species that would correspond to the tissue criterion would not be obvious. Since bioaccumulation models require that components of the food web to be described (e.g., dietary composition, lipid fraction, growth rate, etc.), the parameterization of bioaccumulation models (or choice of bioaccumulation factors) would be ambiguous in such cases. Although tissue criteria for aquatic life would be derived within specified assemblages (e.g., vertebrates, invertebrates), the dietary habits and chemical exposure potential of species within these assemblages can vary widely (e.g., from herbivory to piscivory). This variation in chemical exposure potential would appear to introduce considerable uncertainty in the translation of a tissue criterion to concentrations in ambient media or the aquatic food web.

- **Relationship Between Chemical Sensitivity and Risk.** A second issue that surfaced relates to potential for discontinuity between a species exposure potential (as defined by dietary composition, chemical uptake rates, etc.) and its inherent sensitivity to the chemical as defined by tissue concentration-effect values. This is perhaps best considered in the context of a SSD composed of tissue-based toxicity data. Although this SSD represents the sensitivity differences among species based on intrinsic (internal) toxicity, this distribution does not necessarily correspond to the distribution of exposure potential (and risk) experienced by these species in response to a given exposure regime. Therefore, the relative differences in "risk" to a set of species could differ considerably from their relative differences in sensitivity as defined by tissue concentrations. In other words, the most sensitive species on a tissue concentration basis may not be the species at greatest risk due to variation in exposure potential among species.

To address these two issues, the Subcommittee is considering the use of "representative species" for translating a tissue criterion to corresponding concentrations in ambient media and components of the aquatic food web. A set of representative species could be selected for each aquatic life assemblage that would span a range of factors related to chemical exposure potential (e.g., different feeding guilds/trophic position). For translating national tissue criteria at a national scale, a set of representative species could be defined *a priori*. For translations at a regional or site-specific scale, the representative species could be defined using information specific to the region or site. Using region or local information to define the representative species may be particularly useful, for example, if certain feeding guilds of fish (e.g., large piscivores) are not found at a particular location. In considering the use of representative species, the Subcommittee notes the following feature that make this option attractive.

1. **Representative Species Can Readily be Defined.** Data related to defining chemical exposure potential (e.g., diet, growth rate, lipid content) are expected to be much more plentiful than tissue-based toxicity data. In some cases, available toxicity data may not encompass species that are among the highest exposed. Thus, the translation of tissue criteria to media (or food web) concentrations would be done on a consistent basis even when the composition of the toxicological data sets varied.

2. **Addresses Discontinuity Between Risk and Intrinsic Toxicity.** Representative species would be defined according to a range of exposure pathways, feeding guilds, and habitat preferences. This would enable one to address variation in exposure potential (and risk) that can occur as a function of food web composition, chemical properties, and chemical distribution (e.g., disequilibrium) between water and sediment.

3. **Maintains Consistency Between Criteria Methods.** The same set of representative species could be used for both the deterministic and probabilistic-based tissue criteria. This would maintain consistency among the criteria derivation methods and lead to prediction of media concentrations that would be less dependent on the nuances of the tissue concentration-effects dataset.

4. **Facilitates Translation to Concentrations in Food Web.** By using representative species, one could also translate the national tissue criterion into concentrations in the diet of the representative species (e.g., benthic macroinvertebrates for fish). This might facilitate easier monitoring on the basis of tissue concentrations.

5. **Amenable to Adjustment by Site or Region-Specific Attributes.** Representative species could be defined on a local or regional level which could help address site- or region-specific concerns regarding bioaccumulation potential.

It is worth noting that the approach above does not assume that the representative species are the species actually "at risk" near the tissue criterion. Rather, it assumes that species with intrinsic sensitivities at or near the tissue criterion (whose identity is unknown) could have a range of exposure potential as defined by the representative species.

3.5.2 Bioaccumulation

Once representative species have been defined for an assemblage, the next step in translating a tissue criterion to media concentrations would involve estimating bioaccumulation potential of the chemical in relation to the representative species. For estimating bioaccumulation potential, the Subcommittee proposes to use a framework similar that used by EPA to derive National Ambient Water Quality Criteria to protect human health (USEPA 2000; 2003). This methodological framework is based on the use of both empirical (e.g., bioaccumulation factors, biota-sediment accumulation factors) and mechanistically-based methods (e.g., food web bioaccumulation model; Gobas 1993) for characterizing chemical bioaccumulation in the aquatic diet of humans. Appropriate modifications of this methodology would need to be made to address the diet of representative aquatic life species, but the basic framework would still apply. Some of the salient features of this methodology include:

- Use of high quality measured data for characterizing bioaccumulation (e.g., BAFs, BSAFs) are generally preferred over modeled estimates in part because factors such as chemical metabolism by biota are addressed.
- A three-phase partitioning model is used to address the effect of dissolved and particulate organic carbon on the bioavailability of nonionic organic chemicals.
- Lipid normalization is used to address the effect of differences in lipid content that occur across different species. Accounting for chemical partitioning to organic carbon and lipids has been shown to reduce variability in BAFs measured for PCBs in the Fox River and Green Bay (Burkhard et al., 2003).
- BAFs are aggregated separately for organisms in different trophic levels in order to account for biomagnification and broad physiological differences that can affect bioaccumulation.
- A fugacity-based food web model (Gobas 1993) is used to estimate bioaccumulation in absence of measured data and when the effect of chemical metabolism is considered negligible or is not known.

Once bioaccumulation has been estimated for the representative species, translation to water would be accomplished by dividing the tissue criterion by the appropriate bioaccumulation factor for each representative species within each of the assemblages (e.g., aquatic vertebrates, invertebrates, plants).

$$\text{Water Criterion}_{i,j}\ (\text{mg/L}) = \frac{\text{Tissue Criterion}_{j}\ (\text{mg/kg})}{\text{BAF}_{i,j}\ (\text{L/kg})}$$

Where "i,j" = the "i^{th}" representative species for the "j^{th}" assemblage.

For each assemblage, conversion of a tissue criterion to corresponding concentrations in the aquatic food web (e.g., macroinvertebrates, zooplankton, algae) could be conducted using trophic transfer factors (TTFs) defined separately for each representative species.

$$\text{Concentration in Food Web Component}_{i,j,k}\ (\text{mg/kg}) = \frac{\text{Tissue Criterion}_{k}\ (\text{mg/kg})}{\text{TTF}_{i,j,k}\ (\text{unitless})}$$

Where "i,j,k" = the "i^{th}" food web component of the "j^{th}" representative species for the "k^{th}" assemblage.

According to this scheme, the end result would be a table of criterion values in environmental media (water, sediment) and applicable components of the aquatic food web (e.g., trophic levels 1, 2, 3, 4, etc.) that would vary according to each representative species defined for that assemblage. An example might look something like a table below, with actual chemical concentrations defined in each of the checked boxes.

	Aquatic Vertebrate Assemblage		
Translated Criterion Concentration	Representative Sp. A (piscivore)	Representative Sp. B (benthic carnivore)	Representative Sp. C (herbivore)
Water	✓	✓	✓
Sediment	✓	✓	✓
Algae/Macrophytes	✓	✓	✓
Zooplankton	✓	✓	
Macroinvertebrates	✓	✓	
Forage fish	✓		

The Subcommittee has not discussed <u>if (or how)</u> a final set of "default" criteria concentrations would be selected among the various possible values using the approach outlined above.

3.6 Thoughts on Site-Specific Criteria

Perhaps the most appropriate opportunity for adjusting tissue-based criteria to reflect site-specific differences would arise in their translation to media concentrations (summarized above). Representative species and bioaccumulation could be defined specifically for the site(s) of concern using site data. It is also conceivable that the specific composition of species used to derive the tissue criterion could be modified to more accurately reflect the occurrence of species at a particular site. Such a procedure (called the "recalculation procedure") currently exists for aquatic life criteria derived using the 1985 Guidelines. In using this approach, it would be important to demonstrate that data for any species that would be eliminated from the national tissue criterion database was <u>not</u> a reasonable surrogate for species occurring at the site. The Subcommittee plans to discuss methods for deriving site-specific aquatic life criteria in the near future.

4 Process for Deriving National Tissue-based Wildlife Criteria

4.1 Background on Development of Wildlife Criteria

Although aquatic-dependent wildlife may be protected by aquatic-life criteria, the procedures do not systematically incorporate information on the toxicological sensitivity or the unique exposure scenarios of wildlife species. In 1987, the Government Accounting Office (GAO) issued a report entitled "National Refuge Contamination is Difficult to Confirm and Clean Up" that documented the contaminant clean-up activities at the Kesterson National Wildlife Refuge and other refuges and the limited federal efforts to develop water quality criteria to protect wildlife and their habitats from the adverse effects of chemical contamination (USEPA 1989). The GAO report recognized that cleaning up contaminated sites is difficult when there is a lack of water quality criteria to determine when wildlife are threatened. EPA agreed to modify the criterion for selenium to include wildlife effects.

A workshop in 1988 entitled "Water Quality Criteria to Protect Wildlife Resources" co-chaired by EPA and USFWS focused on evaluating the need for wildlife criteria and developing a strategy for producing wildlife criteria (USEPA 1989). The recommendations from the workshop were that 1) the process for ambient water quality criteria should be modified to consider effects on aquatic-dependent wildlife and 2) chemicals should be prioritized based on their potential to adversely impact wildlife species. In 1989, a preliminary chemical screening was conducted to 1) evaluate whether existing water quality criteria would be protective of wildlife and 2) prioritize chemicals for their potential to adversely impact wildlife species. The approach for screening was derived from an approached developed by the State of Wisconsin for deriving criteria to protect wildlife and domestic animals, which was derived from non-cancer human health criteria. The screening approach considered toxicity and bioconcentration assuming oral ingestion via food and water consumption. The screening study identified the following classes of chemicals for which current water quality criteria may not be adequate to protect wildlife: chlorinated alkanes, chlorinated benzenes and chlorinated phenols, metals, dioxins, and DDT.

The EPA refined this approach in 1991 in an internal report developing interim wildlife criteria. The objective of this analysis was to assess the validity of the previous screening exercises and to evaluate the availability of high quality wildlife toxicity data for criteria development. The report identified chemicals where the interim wildlife criterion was lower than the aquatic life and human health criteria. It also acknowledged that generation of additional wildlife criteria will be difficult due to the lack of toxicity data.

The same approach to wildlife criteria development was being developed at the same time through collaboration with the Great Lakes Water Quality Initiative (GLWQI) for deriving criteria for protection of wildlife species in the Great Lakes. The basic approach used the following model for calculating a wildlife value expressed as the water concentration of a contaminant that, if not exceeded, should be protective of wildlife populations:

$$WV\ (mg/L) = \frac{TD\ (mg/kg\ bw/d) * (1/(UF_A * UF_S * UF_L)) * BW\ (kg)}{W\ (L/d) + \sum[FC_i\ (kg\ food/d) * BAF_i\ (L/kg)]}$$

where:

WV = wildlife value
TD = toxic daily dose
UF = uncertainty factors for interspecies variation (UF_A), subchronic to chronic (UF_S), and LOAEL to NOAEL (UF_L)
BW = body weight of species of concern
W = amount of daily water consumption
FC_i = amount of daily food consumption from the i^{th} trophic level
BAF_i = bioaccumulation factor for the i^{th} trophic level

The toxic daily dose (TD) was derived from an assessment of available toxicity data for a specific chemical. Historically, it was based on an endpoint from the study judged to represent the strongest scientific quality and highest relevance to the assessment. Typically, the TD is calculated using the no-observed-adverse-effect level (NOAEL) or, if necessary, the lowest-observed-adverse-effect level (LOAEL) converted to a daily dose (mg/kg body wt/day). Uncertainty factors are applied to address variation in species sensitivity to the chemical (i.e., UF_A) and deficiencies in study design (i.e., UF_S and UF_L).

Several GLWQI reports related to the process for developing wildlife criteria were published in 1995, including a report detailing the calculation of wildlife criteria for DDT, mercury, 2,3,7,8-TCDD, and PCBs (USEPA 1995c) and a technical support document that presented the rationale for the approach (USEPA 1995d). The wildlife value was calculated based on both drinking water and dietary routes of exposure and was expressed as the chemical concentration in water that would be protective of wildlife. Wildlife values were calculated for several bird and mammal species chosen to represent the Great Lakes aquatic-dependent wildlife, with the final wildlife value for each taxonomic class based on the geometric means of species-specific values.

While work on wildlife criteria for the Great Lakes was nearing completion, work continued on developing approaches for use in developing national wildlife criteria. The EPA Science Advisory Board (SAB) held a meeting in April of 1994 to review progress on development of a national wildlife criteria program. Their primary recommendations were that the program should 1) be guided by the agency's Ecological Risk Assessment Framework, 2) develop a national methodology that can be used to derive regional or site-specific wildlife criteria, 3) use case studies to validate models and methodologies, and 4) focus of protection of wildlife populations, as opposed to individuals (USEPA Science Advisory Board 1994).

The Mercury Study Report to Congress (USEPA 1997) used the GLWQI approach, with a few minor modifications, to develop a national wildlife value for methylmercury in water protective of birds and mammals. The modifications primarily involved changes in the list of representative species and estimates of their diets and the use of additional information to reinterpret toxicity information and the use of uncertainty factors. The report also demonstrated how the approach could be used to calculate the chemical concentration in dietary components representing various trophic levels.

The Canadian government developed a similar approach for national wildlife criteria, with a few notable differences compared to the GLWQI approach (CCME 1998). First, instead of using NOAEL values as the test dose, they use the geometric mean of NOAEL and LOAEL. Second, instead of basing a class-specific criterion on the geometric mean of the wildlife values for representative species, they use the lowest wildlife value calculated for a list of 28 avian or 9 mammalian species to calculate class-specific reference concentrations. Third, the reference concentrations are expressed as the chemical concentration in the diet of each representative species, though the approach does not address the relationship in concentrations among dietary items from the various trophic levels. Wildlife reference values have been developed for DDT, methylmercury, toxaphene, PCBs, and polychlorinated dibenzo-*p*-dioxins/polychlorinated dibenzofurans.

Wildlife criteria also have been developed for specific locations to address questions about the risks of waterborne contaminants to wildlife species. In New Jersey, the GLWQI approach is being used for developing wildlife criteria for PCBs, DDT and mercury for addressing concerns for bald eagles (*Haliaeetus leucocephalus*) and peregrine falcons (*Falco peregrinus*) related to the Endangered Species Act of 1973 (Buchanan et al., 2001). A modification of the GLWQI approach was used in California for addressing a question about the extent to which the EPA proposed human health criterion for methylmercury in the diet was protective of the state's threatened and endangered species (USFWS 2003). The approach was modified to calculate a wildlife value based on dietary concentration of methyl mercury and to convert that value to the corresponding concentrations in dietary constituents from the various trophic levels.

4.2 Issues in Developing Tissue-based Criteria Protective of Wildlife

The process for developing tissue-based criteria for aquatic-dependent wildlife is conceptually the same as for aquatic organisms. However, there are several specific differences in methods that reflect differences in chemical exposure pathways for wildlife and the nature and availability of wildlife toxicity testing. Also, the process for wildlife criteria being discussed is conceptually the same as has been used in previous development of wildlife criteria, although we are revising or expanding some aspects of the process.

The next several sections will discuss the process envisioned for national-level wildlife criteria. This process also is intended to be flexible for modifying a criterion to incorporate site-specific information. Although our current focus has been on an overall national process, guidance will be provided in future versions for determining when and how a site-specific criterion may be developed.

The following sections are intended to describe the wildlife criteria process conceptually, rather than to provide detailed procedures with supporting technical documents. In describing the process, we will highlight those aspects that differ from previous uses of wildlife criteria. One of the primary changes compared to previous wildlife criteria efforts is that we are proposing that a criterion may be based on either a tissue concentration in species of concern or their diet, depending on the availability and quality of information. Another change is to promote the use of probabilistic methods in formulating a criterion when data are of appropriate quantity and quality to do so. For most chemicals with limited data it is recognized that deterministic methods may be more appropriate.

4.3 Screening Available Toxicity Data

The first step in determining how to proceed with a particular chemical is to screen the available toxicity information and collect studies that satisfy the minimum standards for acceptability. The term "study" may refer to a single experiment (or similar unit of research) that estimates a toxicological effect level of a species or a series of experiments that can be integrated to estimate an effect level for a species. Both laboratory and field studies are to be considered if they meet all of the following standards:

- Studies must be based on an experimental design or approach that provides a defensible, chemical-specific response on endpoints that could have implications at the population level, such as reproductive or developmental success, organism viability or growth, etc. For instance, a study must have suitable controls or reference conditions.

- Laboratory studies must contain sufficient information so that the form of the chemical tested is clearly stated, and the administered doses are either reported or can be calculated from information provided.

- Studies must include a subchronic or chronic exposure duration. Laboratory acute oral (i.e., LD50) and short-term (e.g., 5-day LC50) tests are not acceptable.

- Laboratory studies should be based on an oral route of exposure. Laboratory studies using non-oral routes of exposure (e.g., intravenous or subcutaneous injections, implants, etc.) are not acceptable. A possible exception is the use of egg injection studies when there is sufficient understanding of the comparable toxicity from maternally-transferred concentrations.

- Studies may be based on effects relative to a dietary exposure concentration OR to a tissue concentration (e.g., egg or liver concentration vs effect), where scientifically justified.

- Studies must exist in a written form that is available to the public (e.g., journal articles, book chapters, published reports) and that either have gone through a defined technical peer-review process or exist in sufficient detail that a technical quality review can be conducted prior to acceptance.

After reviewing the available studies for a specific chemical, if no study satisfies these standards for a particular taxonomic assemblage (i.e., three wildlife assemblages are birds, mammals, and reptiles), there is insufficient toxicity information to establish a wildlife value for that assemblage. If one or more studies satisfy the standards, they are further evaluated to determine the quality of the study and to document the species tested, endpoints measured, and how endpoints are expressed. This review provides a compilation of all toxicity information for making a preliminary assessment of the quality and quantity of data available for supporting different forms of criteria. It is important to determine how many species from each assemblage have been tested, the comparability of experimental designs and endpoints measured, and any deficiencies in designs that may be addressed through the use of uncertainty factors.

Where multiple studies for a chemical satisfy the standards, the studies need to be evaluated to determine if the quality of information is sufficient to calculate a species sensitivity distribution (SSD). If sufficient toxicity data exist to estimate a representative SSD, the toxicity value used in calculating a criterion would be selected from the distribution (e.g., 5^{th} or 10^{th} percentile from SSD) depending on the intended level of protection. If there is insufficient information for defining a SSD, studies are evaluated to select the one study (or series of related studies) for each taxonomic group that represents the most complete, scientifically-sound study on which to define the test dose (i.e., TD) or tissue concentration for use in a deterministic criterion.

Since a tissue-based wildlife criterion may be based either on chemical concentrations in specific animal tissues (e.g., mg/kg tissue) or in dietary items (e.g., mg/kg food type), the review of toxicity information will group relationships based on animal tissues separately from those based on diet and proceed to derive wildlife values using both types of relationships in parallel. Some studies will provide information on the relationship of effects to both dietary concentrations and tissues concentrations. Ultimately, the weight of the evidence will be used to determine the most scientifically sound means for expressing a wildlife criterion in terms of diet or tissues concentrations.

4.4 Wildlife Criteria Based on Diet Concentration

Wildlife criteria based on dietary concentrations differ somewhat from the tissue-based criteria described for aquatic organisms. The focus on studies that relate effects to chemical concentrations in the diet reflects that this is a commonly used experimental design in wildlife toxicity testing. The core of this approach for wildlife criteria is to determine a daily dietary dose of a chemical that is protective of the more sensitive species of concern and integrate this with information on exposure potential to estimate a concentration in the diet of representative species (also known as a wildlife value) that is intended to be protective (Figure 3). The primary difference in this process based on dietary concentrations compared to the more generalized tissue-based criterion process in Figure 1 is that the dietary toxicity information is integrated with exposure parameters for each representative species prior to calculating the wildlife values. The following sections describe the overall approach for determining a wildlife criterion based on dietary concentrations using deterministic or probabilistic methods.

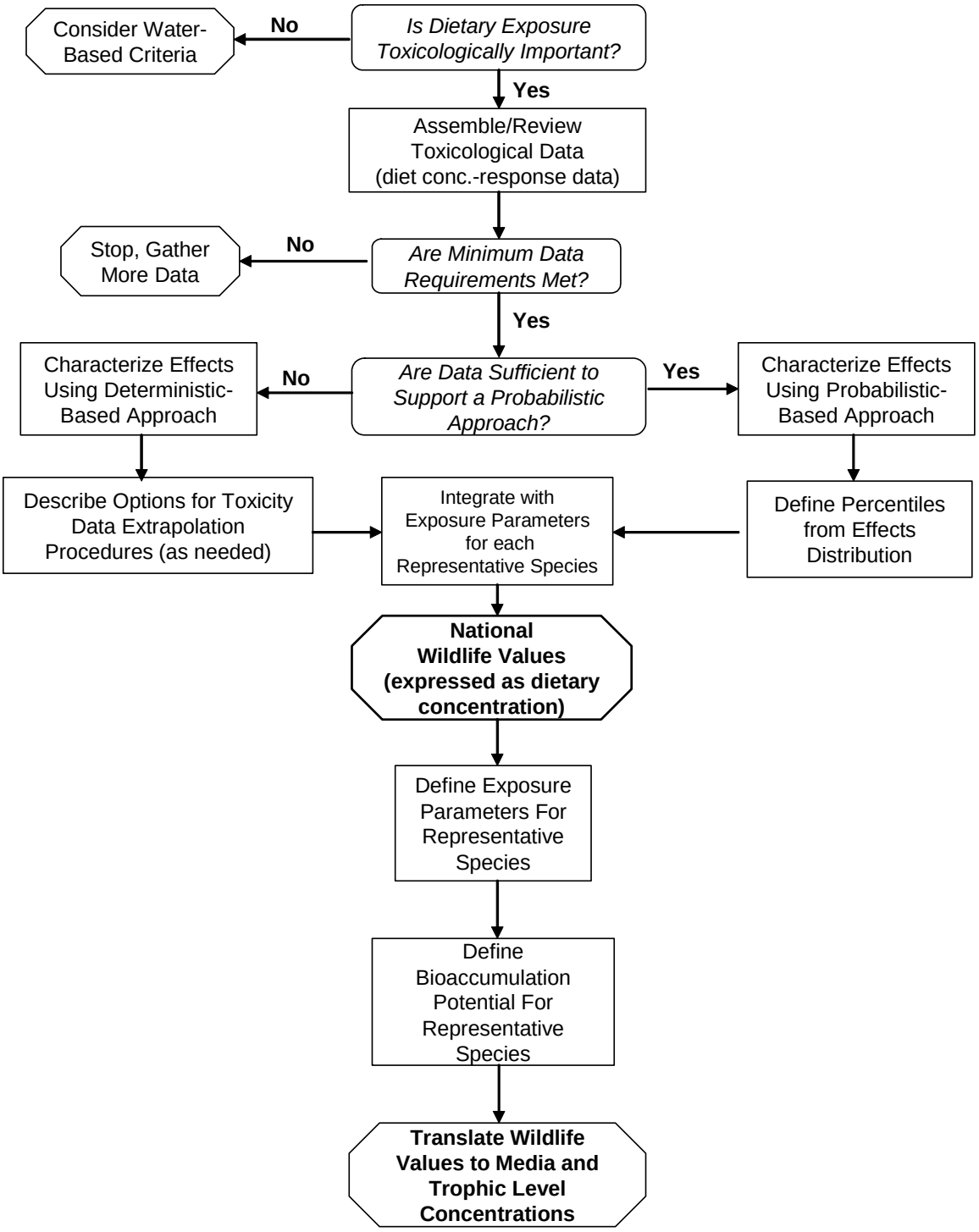


Figure 3. Schematic for Deriving Wildlife Criteria based on Dietary Concentrations

4.4.1 Characterizing effects for a deterministic criterion

For dietary studies that meet the minimum standards of acceptability, endpoints expressed as the chemical concentration in the diet need to be converted to an estimated daily dietary dose (mg/kg body wt/day). The test dose for deterministic criteria should be derived from the study (or series of related studies) for each taxonomic assemblage that represents the most complete, scientifically-sound study. The standards for selecting the most appropriate study are:

- Studies that were designed to measure effects on a suite of reproductive and/or developmental endpoints, as well as record effects on survival, are preferred over studies that are not designed to address reproductive effects.

- Integrative reproductive endpoints that most closely reflect measures of annual fecundity rates are preferred over reproductive endpoints reflecting specific aspects of the reproductive process. For example, while all endpoints may be useful, the number of fledglings produced per nesting attempt is preferred over endpoints such as eggshell quality or number of eggs laid.

- In general, the exposure duration of studies should result in the maximum severity of effects. For bioaccumulative chemicals, studies using a chronic exposure duration are preferred over studies with shorter durations (i.e., subchronic) which may underestimate the severity of effects observed at chronic exposures.

- Studies that clearly define the concentration or dose below which adverse effects are not observed (e.g., NOAELs, EC_{xx} from regression analyses) are preferred over studies where either all of the concentrations or doses cause significant adverse effects (i.e. provide an unbounded LOAEL) or none produce effects distinguishable from control responses (i.e., unbounded NOAEL).

- Laboratory studies that are designed to address the relationship of their results to field responses are preferred over laboratory studies that do not address the relationship to field responses.

- Studies using aquatic-dependent wildlife species would be preferred over studies using species that do not forage on aquatic organisms.

- Field studies that meet the above criteria would be preferred over laboratory studies assuming that relationship between exposure and effects can be accurately described. For populations exposed to multiple chemicals, an explanation is required addressing how observed effects can be assigned to the chemical of concern.

- If more than one study satisfies all of these standards, the study with the highest statistical power would be preferred.

Once the most complete study for selecting a test dose (TD) is determined for each assemblage, it needs to be evaluated to determine if it reflects a daily dietary dose that is protective of the more sensitive species within the assemblage. If not, the selected TD may need to be modified by uncertainty factors. There are two types of uncertainty factors. First, an interspecies uncertainty factor (UF_A) can be used address the uncertainty concerning the variability in toxicological sensitivity among species. Second, there are uncertainty factors used to compensate for deficiencies in the experimental designs of selected studies such as studies of insufficient duration (i.e., subchronic to chronic uncertainty factor or UF_S) or that do not estimate an effects threshold (i.e., LOAEL to NOAEL uncertainty factor or UF_L). By selecting the most complete study for determining the TD, we are striving to minimize the use of the second kind of factors.

During the GLWQI a technical basis for the use of uncertainty factors was developed from an analysis to toxicity studies (USEPA 1996). While this provides an empirical basis on which uncertainty factors can be parameterized, the selection of numerical uncertainty factors also is based on the management goals for the intended level of protection.

The selected study may not provide the most sensitive response among the studies available (i.e., may not result in the lowest TD). Even if the study selected provided the lowest TD among available studies, it is probable that there are additional untested species that are more sensitive to the chemical. Differences in toxicity among wildlife species can often exceed two orders of magnitude (Hart et al. 2001). Analyses of wildlife toxicity databases, primarily acute toxicity test data, have produced several statistical procedures for deriving extrapolation factors for estimating the sensitivity of untested species (Baril et al. 1994, Luttik and Aldenberg 1997, Mineau et al. 1996, 2001a). Although there is much less data to conduct similar analyses of reproduction data, Mineau et al. (2001b) considered that avian reproductive data would be at least as variable as acute toxicity. Luttik et al. (2005) review these methods and propose an approach for extrapolating long-term toxicity data among wildlife species. These methods provide an empirical basis for estimating a UF_A for use in deterministic criteria where there are insufficient data to calculate a chemical-specific SSD.

In order to develop a criterion that is protective of more sensitive aquatic-dependent species in a taxonomic class, the TD from the selected study is modified using a UF_A that integrates information available from the empirically-based methods above together with the toxicity data available from all species in studies evaluated under the second set of criteria. Although the empirically-based methods provide insights into the variation in sensitivity among species based on analysis of large datasets, a comparison of the TD from the selected study with endpoints from the other available studies provides insights into where the selected TD falls within the distribution.

The Subcommittee recognizes that for some chemicals the most complete study available may be of insufficient duration or may not clearly define an effects threshold and uncertainty factors may be appropriate to address these deficiencies. The guidance developed during the GLWQI on the use of uncertainty factors provides a basis for developing national-level wildlife criteria (USEPA 1996), but we have not fully discussed what modifications or additions are needed.

In characterizing the effects information for a deterministic wildlife criterion, the selected TD is divided by the product of the three uncertainty factors. Concern has been expressed that multiplying several uncertainty factors can result in criteria that are overprotective. In the process outlined above, it is recognized that the UF_A may be important to protect species that are thought to be more sensitive than the tested species, but the UF_L and UF_S are used to compensate for deficiencies in the experimental designs of available studies. The use of a UF_A may be appropriate for most chemicals unless there is evidence that the test species is also a relatively sensitive species with its taxonomic assemblage. On the other hand, studies requiring the use of UF_L and UF_S should be used only when no other studies are available. For chemicals where the only studies available would require use of both a UF_L and UF_S, the uncertainty in toxicity information may be so great that no criterion should be established.

We have additional work to do in providing guidance on when it is appropriate to consider uncertainty factors and how to parameterize them. We are striving to develop a process that minimizes the need for uncertainty factors, and when they are needed, to provide guidance for determining an empirically-based value or concluding that the uncertainty is too great for criterion development.

4.4.2 Selecting toxicity information for a probabilistic criterion

When it is determined that there is a sufficient number of studies with different species in order to calculate an SSD, the studies are evaluated further to ensure that they are of comparable quality. An SSD is most useful if it accurately reflects the difference in chemical sensitivity among species. If studies vary too much in the endpoints measured, duration of exposure, statistical power, or other experimental design features, the calculated SSD may be confounded by these experimental differences that mask the true differences in sensitivity. It is also possible that some studies use exposure concentrations that do not result in a fully described dose-response relationship or identification of an effects threshold (i.e., only produce an unbounded LOAEL). In such cases, limited use of UF_L may be warranted to keep an adequate sample size of species tested. Similarly, for studies that are considered to be of insufficient exposure duration, it may be warranted to use UF_S to estimate what an effects threshold would be under chronic exposure scenarios. However, the more studies that need to be amended by UF_S or UF_L to compensate for deficiencies, the greater the uncertainty that the resulting SSD is an adequate reflection of the distribution of species sensitivities.

Criteria also can be expressed probabilistically when one or more studies quantify a dose-response relationship for a population-relevant endpoint. Instead of relying only on an estimate of an effects threshold, such as an NOAEL, a criterion derived using a dose-response relationship can be presented as an equation that estimates the exposure concentration associated with any level of effect.

The Subcommittee has not yet discussed the quantity or quality of data required to consider the use of probabilistic methods for criteria based on dietary concentrations.

4.4.3 Characterization of Exposure for National Criteria

Although there may be limited toxicological information for estimating the sensitivity of a specific wildlife species or the range of sensitivities among aquatic-dependent wildlife species,

the variation in exposure potential among aquatic-dependent wildlife species can be estimated where there is information about the dietary composition of species, their food consumptions rates (either measured or estimated as a function of body weight), and the relationship of chemical concentrations among various trophic levels of dietary items (i.e., trophic transfer factors). Aquatic-dependent wildlife species vary greatly in their dietary composition (e.g., aquatic vegetation vs higher trophic level fish, entirely aquatic diet vs partially aquatic diet), which results in great variation in exposure potential. The dietary composition of some wildlife species also can vary geographically due to differences in prey availability. Body size is important because food consumption rates tend to increase with decreasing body size. Also, trophic transfer factors vary depending on the bioaccumulation characteristics of a chemical, which affects the exposure potential among species.

For national-level criteria, trophic levels will be defined similarly to those used in the GLWQI assessment, with trophic level 1 (i.e., TL1) representing primary producers, TL2 representing primary consumers (i.e., many invertebrates and small fish), TL3 representing secondary consumers (e.g., forage fish, insectivorous birds), and TL4 representing top predators (e.g., carnivorous fish, fish-eating birds).

The wildlife criteria process is designed to determine which species have high exposure potential based on the factors above.

4.4.4 Representative species

There is a large number of wildlife species whose diet is derived entirely or partially from aquatic foodwebs. Many assessments will not estimate the exposure potential for every aquatic-dependent species, but will select a subset of species to represent the diversity of factors that determine exposure potential. Wildlife species identified as “representative species” are not necessarily the species of greatest concern or the only species being considered, but are chosen to represent the range of aquatic-dependent species. In other words, just because a selected representative species does not inhabit a certain location does not means it is not representing similar species that do. For each chemical the process is designed to identify which foraging strategies have high exposure potential.

For national-level wildlife criteria, a table of representative species is being developed that 1) reflects the diversity in body weights and diets among aquatic–dependent species, 2) includes species that have been studied sufficiently to quantify dietary composition and determine trophic level of dietary components, and 3) are relatively widely distributed and recognized and/or valued by the public. At a later stage, guidance will be provided for implementers that prefer to select species representative of their specific jurisdiction, including issues to address in providing a rationale for their selection.

A dietary composition and trophic level analysis was completed for 20 species for the GLWQI (USEPA 2002). The list includes 16 birds (including Osprey, Bald eagle, Belted kingfisher, Herring gull, Ring-billed gull, Great blue heron, Black-crowned night-heron, Common tern, Forster's tern, Caspian tern, Double-crested cormorant, Common merganser, American merganser, Red-breasted merganser, Lesser scaup, and Mallard,) and four mammals (including Mink, River otter, Raccoon, and Harbor seal). Work is currently underway to expand this analysis to additional species. Candidate species under consideration include Common loon,

Western grebe, Pied-billed grebe, Eared grebe, White pelican, Green heron, Little blue heron, King rail, Peregrine falcon, Least tern, American avocet, and Marsh wren. We have decided to initially focus on species feeding primarily in freshwater systems, though in the future the process will be expanded to address species feeding in marine and estuarine systems. We also have not decided how to address reptiles given the paucity of toxicological data.

4.4.5 Body weight, food ingestion rate, and diet composition of representative species for deterministic criteria

Smaller birds and mammals generally have higher food ingestion rates relative to their body mass than do larger ones. This suggests that small animals would be exposed to a larger quantity of contaminants relative to the body size (i.e., dose) than larger animals. However, small piscivores are generally size-limited predators, and feed on smaller fish in lower trophic levels than do larger piscivores. Because the concentration of bioaccumulative chemicals usually is lower in lower trophic level organisms, it is not clear that small animals always experience higher exposures than larger animals. Therefore, to identify species likely to experience the highest exposure levels, both relative food ingestion rates and the trophic level of prey must be considered. For highly bioaccumulative chemicals, the species feeding at highest trophic levels of the aquatic food chain may have the highest dietary exposure potential (i.e., result in lowest criterion) in the process. For chemicals with lower bioaccumulation potential, the smallest body mass (and consequently highest food ingestion rate) may have the highest dietary exposure potential.

For national-level deterministic wildlife criteria, default values representing female body weight, estimated food ingestion rate (FIR), and proportion of diet derived from each trophic level category will be selected for each representative species. The reference for each default body weight and FIR will be stated, as well as the background analysis for determining the trophic level proportions for each diet. Implementers at the state or site-specific level will be able to use locally-derived information for modifying these default values if they can provide a rationale for why that is an improvement over using the national default information.

4.4.6 Trophic transfer factors for deterministic criteria

Trophic transfer factors (TTF) represent the ratio of the estimated chemical concentration in one trophic level to the concentration in the trophic level below it. They may be calculated directly from measured concentrations in representatives from various trophic levels or indirectly from the ratio of BAFs. The Subcommittee has not discussed yet the requirements for determining TTF for a national-level criterion, although we recognize the need for establishing a transparent process.

4.4.7 Body weight, diet composition, and trophic transfer factors for probabilistic criteria

The point estimate for each of the exposure parameters used in a deterministic criterion is derived from a distribution of values. Some of these distributions are well-defined descriptions of the natural variability for a parameter, while we are more uncertain in our knowledge about others. In a probabilistic approach to developing criteria, each of the exposure parameters can be described as a distribution in order to better understand the variability and uncertainty in the exposure potential of individuals within a species. The subcommittee has not specifically discussed procedures for accomplishing this.

4.4.8 Calculating a deterministic wildlife value based on dietary concentration

The model used to calculate a wildlife value based on concentration of chemical in the diet is a modification of the model used in the GLWQI expressed as the concentration in diet rather than water (USEPA 1995b). It is the same as the model used in an assessment of mercury in California (USFWS 2003). A wildlife value is calculated for each representative species because exposure potential varies with body weight, food ingestion rate, and diet composition, even though the same test dose is applied to each representative species (Figure 3). The equation for calculating wildlife values is:

$$WV_{food}\ (mg/kg\ food) = \frac{TD\ (mg/kg\ bw/d) * (1/(UF_A * UF_S * UF_L)) * BW\ (kg)}{\sum[FC_i\ (kg\ food/d)]}$$

where:

WV = wildlife value expressed as the chemical concentration in the diet of each representative species,
TD = test dose expressed as daily dietary dose from selected study,
UF = uncertainty factors for interspecies variation (UF_A), subchronic to chronic (UF_S), and LOAEL to NOAEL (UF_L),
BW = estimated mean body weight of a representative species, and
FC_i = amount of daily food consumption for each species from the i^{th} trophic level.

Because it is based on the entire diet for each species, a wildlife value itself is not a criterion. The wildlife value of each species needs to be translated into the corresponding concentrations at each trophic level using the estimates for the amount of food consumed from each trophic level and the TTFs. For example, a wildlife species that feeds on both TL3 and TL4 fish will have a wildlife value that reflects the concentration in the entire diet which may correspond to a concentration in TL3 fish that is lower than the wildlife value and a concentration in TL4 fish that is higher. This is needed to compare the estimated risk among representative species regardless of diet composition by translating the wildlife value into a common currency, such the corresponding concentration in TL3 fish. Because of significant differences in diet composition among species, the species with the lowest wildlife value does not necessarily translate into the species with the lowest corresponding concentration in TL3 fish.

A table will be produced listing the wildlife value for each representative species within the bird and mammal assemblages and the corresponding concentrations in the various trophic levels and water that would result in an average dietary concentration equivalent to the wildlife value. This table would show how differences in diet composition and body weight among species influence the chemical concentrations in each trophic level considered to be protective. It should be remembered that the same daily dietary dose is used for each species within an assemblage in these calculations assuming that any one of these untested species could be among the more sensitive species in the assemblage distribution. Consequently, the table does not literally specify which species are at greatest risk, but it does indicate which types of exposure profiles may be at greatest risk if species with those exposure profiles are among the more sensitive species toxicologically.

The information presented in these tables will provide the basis for determining national wildlife criteria, but more detailed procedures for determining criteria values will not be decided until discussions with Office of Water management have occurred. The final criterion will reflect the management goals for the intended level of protection.

4.4.9 Calculating probabilistic wildlife values based on dietary concentration

There are a variety of ways that wildlife values could be expressed probabilistically. One or more of the parameters in the equation above could be expressed as a distribution or an equation. This could result in wildlife values describing a distribution of values (rather than a single value) or the probability of exceeding a specified value. The goal should be to improve the characterization of risks by more explicitly integrating natural variability and uncertainty into the calculation of wildlife values. This provides risk managers greater insight in the degree of uncertainty in calculating wildlife values and the ramifications for achieving the intended level of protection.

The Subcommittee has not yet discussed specific applications of probabilistic methods or the data needs for using these methods.

4.5 Wildlife Criteria Based on Tissue Concentrations

Wildlife criteria based on tissue concentrations are quite similar to the tissue-based criteria described for aquatic organisms. They focus on studies from the laboratory or field that relate effects to chemical concentrations in specific animal tissues. Beyer et al. (1996) reviewed the evidence for interpreting tissue concentrations in wildlife. The core of this approach for wildlife criteria is to determine a specific tissue concentration of a chemical that is protective of the more sensitive species of concern (Figure 4). The following sections describe the overall approach for determining a wildlife criterion based on tissue concentrations using deterministic or probabilistic methods.

4.5.1 Characterizing effects for a deterministic criterion

In some cases the relationship between a specific animal tissue concentration and population-relevant effects may be less uncertain and more repeatable that relationships between diet concentrations and the same effects. When this can be demonstrated, a wildlife criterion based on specific tissue concentrations may be more robust at defining an unacceptable risk to wildlife species than one based on dietary concentrations (Figure 4). This process would be the same conceptually as was described for aquatic organisms in Section 3 and as illustrated in Figure 1. The information on tissue concentration-response relationships may or may not be derived from studies that also provide diet concentration-response information. Some studies, especially field studies, may provide tissue concentration-response information with little or no information about corresponding dietary exposure. Tissue concentration studies that do not provide direct evidence of the relationship back to dietary exposure may be acceptable if there is an alternative method to estimate corresponding exposure concentrations. A method for translating tissue concentrations into corresponding dietary concentrations is needed in order to understand how the chemical concentrations in tissues relates to concentrations throughout the food web and in abiotic media.

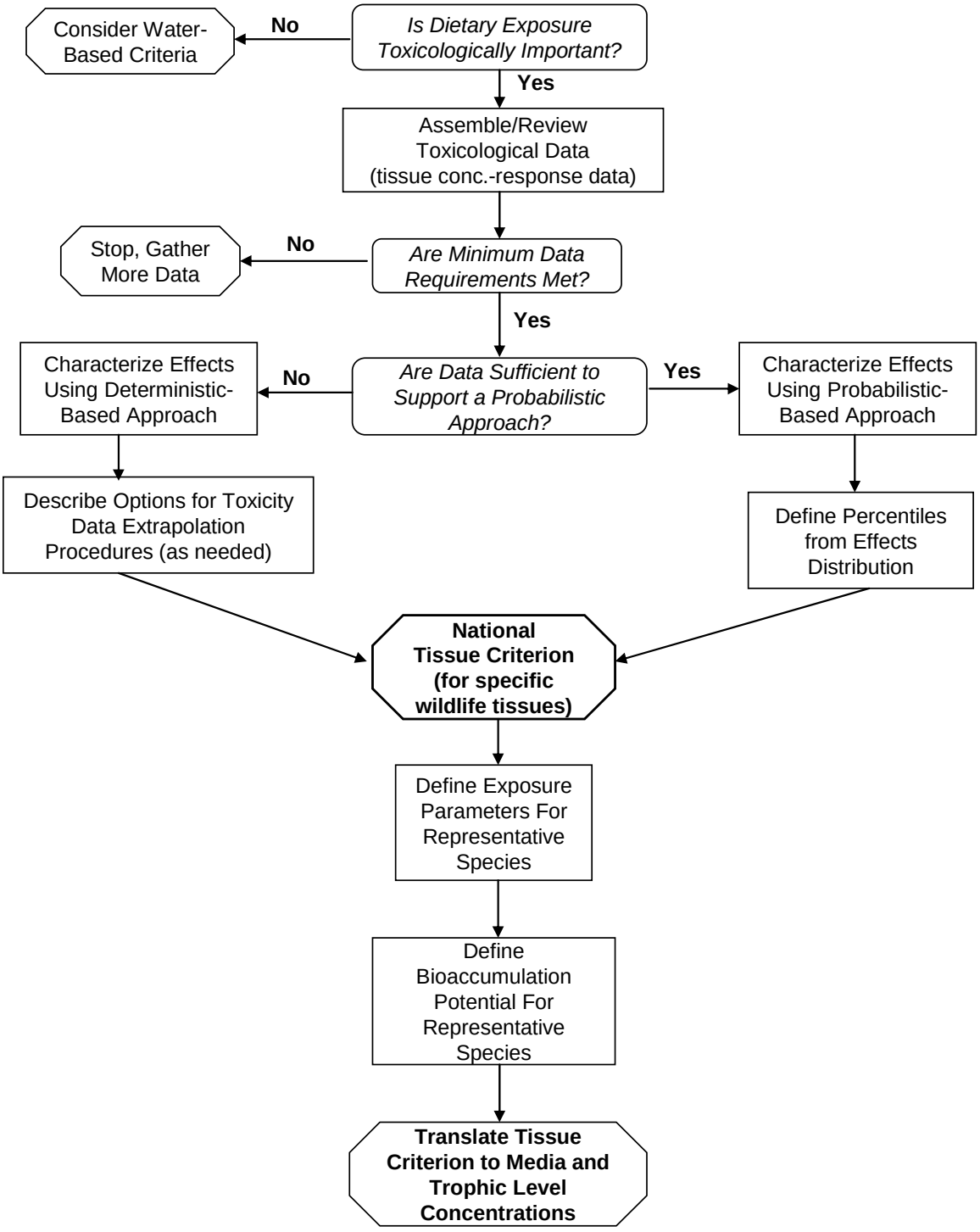


Figure 4. Schematic for Deriving Wildlife Criteria based of Tissue Concentrations

The standards outlined in section 5.4.1 for selecting the most appropriate study for each assemblage also apply here. An additional standard relates to the adequacy of the type of tissue used in the relationship.

- Studies with tissue concentration-response information based on tissue types with a direct causal relationship to the observed effects are preferred over studies based on tissue types with no clear causal connection. For example, if a major effect of a specific chemical is to interfere with embryo development and reduce hatchability, relationships based on whole egg concentrations may be the most appropriate tissue.

4.5.2 Characterizing effects for a probabilistic criterion

For some chemicals, there may be multiple studies that have determined the relationship between effects and the chemical concentrations in the same type of tissue. When it is determined that there is a sufficient number of studies with different species in order to calculate an SSD, the studies are evaluated further to ensure that they are of comparable quality. An SSD for tissue concentrations would provide additional information for determining a protective concentration commensurate with management goals.

The Subcommittee has not yet discussed the quantity or quality of data required to consider the use of probabilistic methods for criteria based on tissue concentrations.

4.5.3 Calculation of a wildlife value based on tissue concentration

Deterministic wildlife values based on tissue concentrations are calculated using the chemical concentration from the selected study modified by uncertainty factors, where necessary. Unlike the wildlife values based on dietary concentrations, the values based on tissue concentrations do not vary among representative species as a function of body weight and food ingestion rate. The equation for calculating wildlife values is:

$$WV_{tissue}\ (mg/kg\ tissue) = TC\ (mg/kg\ tissue) * (1/(UF_A * UF_S * UF_L))$$

where:

WV = wildlife value expressed as the chemical concentration in the specified tissue,
TC = tissue concentration from selected study, and
UF = uncertainty factors for interspecies variation (UF_A), subchronic to chronic (UF_S), and LOAEL to NOAEL (UF_L).

The uncertainty factors used to modify tissue concentrations are conceptually the same as those used for dietary concentrations. However, the empirical relationships based on tissue concentrations and effects may differ from those examined for dietary concentrations in the GLWQI guidance for use of uncertainty factors (USEPA 1996). Consequently, it should not be assumed that uncertainty factors selected for dietary studies apply equally well to tissue concentration information. Justification for the use of uncertainty factors should be based on a separate analysis of existing data. The Subcommittee has not specifically discussed the approach to using uncertainty factors for wildlife values based on tissue concentrations.

The wildlife values based on tissue concentrations may be used directly in determining a wildlife criterion. Additionally, there needs to be a method for translating these tissue concentrations into corresponding concentrations in lower trophic levels and abiotic media. This is straightforward when the relationship between effects and both diet and tissue concentrations are derived from the same study. When these relationships are based on different studies, we need to be able to compare wildlife values based on diet vs tissue concentrations to understand the relative degree of protection afforded. The Subcommittee is currently conducting an empirical analysis of the relationships between effects and concentrations in both diet and tissues to support methods to use in the criteria development process. In the future, guidance will be developed for determining wildlife criteria when there is sufficient data to calculate wildlife values based on both diet and tissue concentrations. The Saginaw Bay mink study series represents a good example of integrating both diet- and tissue-based data sets (Tillitt et al. 1996).

4.6 Role of population modeling

Under certain circumstances population modeling could be a valuable tool in the development of tissue-based wildlife criteria as a means of understanding the consequences of chemical exposures to wildlife populations. The primary use of population modeling would be in the development of site-specific criteria for data-rich chemicals where it is possible to improve the characterization of risks beyond that possible with the methods above. However, we do not envision the use of population modeling in national- or regional-level wildlife criteria development along the lines of that articulated by the Water-based Criteria Subcommittee (WCS). There are several important reasons for this difference in approach.

First, the WCS is using population models as a means of integrating effects data on survival, growth, and reproduction into a common metric of change in population size. However, the bulk of wildlife chronic effects data for bioaccumulative chemicals relates to reproduction endpoints, with little or no data on effects of chronic exposure to survival rates. Also, for many chemicals the effects to reproductive endpoints typically occur at environmental concentrations that are lower than would be expected to affect survival. Consequently, there is not the same possibility or need for integrating survival and reproduction data.

Second, the WCS is using population models to integrate population responses to varying exposure concentrations over time where there are periods of exposure causing declines in a population and periods of recovery. Wildlife exposed to bioaccumulative chemicals through an aquatic food web are expected to have less variation in exposure concentrations over time, and criteria are based on exposure concentrations deemed acceptable over the long-term. Wildlife criteria are intended to prevent the types of population-level effects that would require a recovery. Consequently, there is not the same need to integrate the effects of variable exposure or consider recovery rates.

Third, population models conceptually can be used to estimate the magnitude of effects to individuals that can be assimilated by a population, leading to estimates of the environmental concentration protective of the population rather than relying on the somewhat lower concentration that would protect against effects to individuals. However, wildlife toxicity information will often come from studies on species other than the species of concern and it will be from studies with less standardization than being required by the WCS. The toxicological

sensitivity of species of concern often has to be estimated from limited data on other species, resulting in an unknown degree of uncertainty in estimates. The amount of uncertainty in estimating the sensitivity of untested species can overshadow attempts to use population modeling to characterize effects in a population context.

Because of these reasons, we concluded that population modeling is unlikely to improve the characterization of risks at the national or regional level beyond what is possible with the deterministic or probabilistic approaches described above. However, we will be discussing the role population modeling could have in improving site-specific criteria development. An important consideration in those discussions is that populations are not exposed to one stressor (chemical or non-chemical) at a time, and population-level assessments in criteria development will need to address the cumulative impacts of co-occurring stressors.

5 Issues for SAB

5.1 Charge Questions

1. For chemicals with a high propensity to bioaccumulate in aquatic food webs and for which diet is a primary route of exposure, the Tissue-based Criteria Subcommittee proposes to develop tissue-based criteria expressed as the chemical concentrations in specific animal tissues or dietary concentrations, with a process for translating to corresponding water and sediment concentrations. Tissue-based criteria allow for integration of multiple exposure pathways (water, diet) and facilitate direct comparison with environmental tissue concentrations to determine if there is a risk of adverse effects. **Please comment on the rationale and conceptual approach used for the development of tissue-based criteria for this group of chemicals. Is the SAB aware of other approaches for deriving criteria for these bioaccumulative chemicals that EPA should consider?**

2. The proposed process for Tissue-based Criteria is intended to be flexible to maximize the use of available data and to accommodate certain limitations in the quality and quantity of data. National-level criteria may use deterministic approaches to characterize toxicity data when data are limited or probabilistic approaches (e.g., species sensitivity distributions) when data are sufficient. The process will also describe how a criterion may be refined on a site-specific basis when additional data are available. **Considering the strengths and limitations of the more flexible approach used to derive tissue-based criteria, please comment on the rationale and preference for allowing flexibility in the procedures used.**

3. Unlike the dynamic exposure scenarios being addressed by the Water-based Criteria Subcommittee, the Tissue-based Criteria Subcommittee is considering a steady-state approach for developing national criteria for bioaccumulative chemicals (i.e., modeling bioaccumulation and toxicity as a function of constant concentrations). Rationale for this approach is due in part to the much slower accumulation kinetics generally associated with these chemicals in higher trophic level fish and aquatic-dependent wildlife and concerns over their long-term bioaccumulation. In the context of population modeling, there appears to be much less residue-response information available for integrating responses of various demographic parameters over multiple life stages, such as fecundity and adult, juvenile, and larval survival. Consequently, the feasibility and utility of integrating population modeling into national-level tissue criteria for bioaccumulative chemicals is not clear to the Tissue-based Criteria Subcommittee. Current thinking is that where sufficient data exist to characterize exposure, bioaccumulation and toxicity on a dynamic basis, population modeling may evolve into an important tool in the development of site-specific criteria. **Please comment on the rationale used by the Tissue-based Criteria Subcommittee for determining if/when to use population modeling in the development of Tissue-Based Criteria.**

5.2 Additional Technical Issues: Aquatic Life Criteria

1. **Toxicity Data Extrapolations.** For deterministic-based aquatic life criteria, which might be derived with as few as 4-5 species per assemblage, the Subcommittee is considering the need for toxicity data extrapolations to account for potentially greater sensitivity of untested species. While methods have been developed conducting interspecies extrapolations of toxicity using water concentration-based toxicity data (e.g., ICE, adjustment factors for secondary Tier II criteria under the Great Lakes Initiative), the Subcommittee is not aware of analogous methods using tissue concentration-based toxicity data. **Please comment on: (1) the need for such toxicity data extrapolations and (2) available methods for conducting such extrapolations using tissue concentration-based toxicity data that the Subcommittee should consider.**

2. **Representative Species.** In order to address differential exposure potential among aquatic species and implementation of tissue-based criteria, the Subcommittee is considering the use of "representative species" (in conjunction with bioaccumulation methods) as described in Section 3.5.1. These representative species would reflect a range of exposure potential that might be experienced by aquatic species with tissue-based sensitivities at or near the tissue criterion. **Please comment on: (1) the rationale and approach presented by the Subcommittee for using representative species and (2) other methods the Subcommittee should consider for translating a tissue criterion into corresponding concentrations in media and the food web.**

3. **Bioaccumulation.** In the revision of the Aquatic Life Criteria guidelines to better address "bioaccumulative" chemicals, the Subcommittee proposes to use a framework for assessing bioaccumulation potential that is similar to that used by EPA in its National Ambient Water Quality Criteria to protect human health (USEPA 2000, 2003). The mechanistically-based portion of the bioaccumulation framework uses the fugacity based food web model of Gobas (1993) in cases where measured data is absent and when metabolism is considered negligible. Food web models have continued to evolve and improve since the publication of the Gobas 1993 model. **Have improvements in these models been significant enough to warrant EPA adopting an improved model into the bioaccumulation methodology of the revised guidelines? Do you agree with the idea of reserving the use of dynamic (time varying) bioaccumulation modeling for situations where short-term fluctuations in media concentrations are a concern with sensitive aquatic species?**

5.3 Additional Technical Issues: Wildlife Criteria

1. **Uncertainty Factors.** The standards for selecting wildlife toxicity studies emphasize the need to select the most complete studies in order to limit the need for uncertainty factors that compensate for deficiencies in experimental designs. When uncertainty factors are needed to maintain the desired level of protection, their selection should be based on an analysis of available information. **Based on the proposed procedures for selecting toxicity data, please comment on the rationale for use of uncertainty factors, where needed.**

6 References

Asfaw A, Ellersieck MR, and Mayer FL. 2003. *Interspecies Correlation Estimations (ICE) for Acute Toxicity to Aquatic Organisms and Wildlife. II. User Manual and Software*. EPA/600/R-03/106. U.S. Environmental Protection Agency, National Health and Environmental Effects Research Laboratory, Gulf Ecology Division, Gulf Breeze, FL. 14 p.

Baril A, Jobin B, Mineau P and Collins BT. 1994. *A consideration of interspecies variability in the use of the median lethal dose (LD50) in avian risk assessment.* Technical Report Series No. 216, Canadian Wildlife Service, Ottawa, Canada.

Barron MG, Hansen JA, and Lipton J. 2002. Association between contaminant tissue residues and effects in aquatic organsisms. *Rev. Environ. Contam. Toxicol.* 173:1-37.

Beyer WN, Heinz GH, and Redmon-Norwood AW. 1996. *Environmental Contaminants in Wildlife: Interpreting Tissue Concentrations.* Lewis Publishers CRC Press, Boca Raton, FL. 494 pp.

Buchanan GA, Russell DW, and Thomas DA. 2001. *Derivation of New Jersey-Specific Wildlife Values as Surface Water Quality Criteria for: PCBs, DDT, Mercury*. Unpublished report.

Burkhard L.P. 2003. Factors influencing the design of bioaccumulation factors and biota-sediment accumulation factors field studies. *Environmental Toxicology and Chemistry.* 22:351-360.

Burkhard LP, Endicott DE, Cook PM, Sappington KG, and Winchester EL. 2003. Evaluation of two methods for prediction of bioaccumulation factors. *Environmental Science and Technology*. 37(20):4626-4634.

Cook PM, Carlson AR, Lee H. 1989. *Tissue Residue Approach*. Sediment Classification Methods Compendium, Chapter 7. EPA 823-R-92-006. Office of Water, Washington, D.C.

Cook PM, Erickson RJ, Spehar RL, Bradbury SP, and Ankley GT. 1993. *Interim Report on Data and Methods for Assessment of 2,3,7,8-Tetrachlorodibenzo-p-dioxin Risks to Aquatic Life and Associated Wildlife.* EPA/600/R-93/055. Office of Research and Development, Duluth, MN 167pp.

Di Toro DM, McGrath JA, and Hansen DJ. 2000. Technical basis for narcotic chemicals and polycyclic aromatic hydrocarbon criteria. I. Water And Tissue. *Environmental Toxicology and Chemistry*. 19:1951-1970.

Di Toro DM, and McGrath JA. 2000. Technical basis for narcotic chemicals and polycyclic aromatic hydrocarbon criteria. II. Mixtures And Sediments. *Environmental Toxicology and Chemistry*. 19:1971-1982.

Di Toro DM, Allen HE, Bergman HL, Meyer JS, Paquin PR, and Santore RC. 2001. A biotic ligand model of the acute toxicity of metals I. Technical basis. *Environmental Toxicology and Chemistry* 20:2383-2396.

Escher BI and Hermens JLM. 2002. Modes of action in ecotoxicology: Their role in body burdens, species sensitivity, QSARs, and mixture effects. Environmental Science and Technology. 36(20): 4201-4217.

Gobas FAPC. 1993. A model for predicting the bioaccumulation of hydrophobic organic chemicals in aquatic food webs: Application to Lake Ontario. *Ecological Modeling*. 69:1-17.

Hart A, Balluff D, Barfknecht R, Chapman PF, Hawkes T, Joermann G, Leopold A, and Luttik R. 2001. *Avian effects assessment: A framework for contaminant studies*. SETAC, Pensacola, FL. 193 pp.

Jarvinen AW and Ankley JT. 1999. *Linkage of Effects to Tissue Residues: Development of a Comprehensive Database for Aquatic Organisms Exposed to Inorganic and Organic Chemicals*. SETAC Press, Pensacola, FL, 364 pp.

Könemann H. 1981. Quantitative structure-activity relationships in fish toxicity studies Part 1: Relationship for 50 industrial pollutants. *Toxicology*. 19(3): 209-221.

Landrum PF, Steevens JA, Gossiaux DC, McElroy M, Robinson S, Begnoche L, Chernyak S, and Hickey J. 2004. Time-dependent lethal body residues for the toxicity of pentachlorobenzene to Hyalella azteca. *Environmental Toxicology and Chemistry* 23:1335-1343.

Landrum PF, Steevens JA, McElroy M, Gossiaux DC, Lewis JS, and Robinson SD. 2005. Time-dependent toxicity of dichlorodiphenyldichloroethylene to Hyalella azteca. *Environmental Toxicology and Chemistry* 24(1): 211-218.

Lee J-H, Landrum PF, and Koh C-H. 2002a. Toxicokinetics and time-dependent PAH toxicity in the amphipod Hyalella azteca. *Environmental Science and Technology*. 36: 3124-3130.

Lee J-H, Landrum PF, and Koh C-H. 2002b. Prediction of time-dependent PAH toxicity in Hyalella azteca using a damage assessment model. *Environmental Science and Technology*. 36: 3131-3138.

Legierse KCHM, Verhaar HJM, Vaes WHJ, De Bruijn JHM, and Hermens JLM. 1999. Analysis of time-dependent acute aquatic toxicity of organophosphorus pesticides: The critical target occupation model. *Environmental Science and Technology*. 33:917-925.

Luttik R and Aldenberg T. 1997. Extrapolation factors for small samples of pesticide toxicity data: Special focus on LD50 values for birds and mammals. *Environmental Toxicology and Chemistry* 16: 1785-1788.

Luttik R, Mineau P, and Roelofs W. 2005. A review of interspecies toxicity extrapolation in birds and mammals and a proposal for long-term toxicity data. *Ecotoxicology* (in press).

McCarty LS and Mackay D. 1993. Enhancing ecotoxicological modeling and assessment: body residues and modes of toxic action. *Environmental Science and Technology* 27:1719-1728.

McCarty LS. 1986. The relationship between aquatic toxicity QSARS and bioconcentration for some organic chemicals. *Environmental Toxicology and Chemistry* 5:1071-1080.

Mineau P, Collins BT and Baril A. 1996. On the use of scaling factors to improve interspecies extrapolation of acute toxicity in birds. *Regulatory Toxicology and Pharmacology* 24: 24-29.

Mineau P, Baril A, Collins BT, Duffe JA, Joermann G and Luttik R. 2001a. Pesticide acute toxicity reference values for birds. *Reviews of Environmental Contamination and Toxicology* 170: 13-74.

Mineau P, Hooper M, Elmegaard N, Grau R, Luttik R and Ringer R. 2001b. Case study 5: Foliar insecticide II. In Hart A, Balluff D, Barfknecht R, Chapman PF, Hawkes T, Joermann G, Leopold A, and Luttik R (eds). 2001. *Avian effects assessment: A framework for contaminant studies.* SETAC, Pensacola, FL. 193 pp.

Munns WR, Black DE, Gleason TR, Salomon K, Bengtson D, and Gutjahr-Gobell R. 1997. Evaluation of the effects of dioxin and PCBs on *Fundulus heteroclitus* populations using a modeling approach. *Environmental Toxicology and Chemistry*. 16(5):1074-1081.

Paquin PR, Gorsuch JW, Apte S, Batley GE, Bowles KC, Cambell PGC, Delos CG, Di Toro DM, Dwyer RL, Galvez R, Gensemer RW, Goss GG, Hogstrand C, Janssen CR, McGeer JC, Naddy RB, Playle RC, Santore RC, Schneider U, Stubblefield WA, Wood CM, Wu KB. 2002. The Biotic Ligand Model: a Historical Overview. *Comparative Biochemistry and Physiology Part C* 133:3-35.

Peterson JA and Nebeker AV. 1992. Estimation of waterborne selenium concentrations that are toxicity thresholds for wildlife. *Archives of Environmental Contamination and Toxicology* 23: 154-162.

Rainbow PS. 2002. Trace metal concentrations in aquatic invertebrates: why and so what? *Environmental Pollution* 120:497-507.

Stephan CE, Mount DI, Hansen DJ, Gentile JH, Chapman GA, and Brungs WA. 1985. *Guidelines for Deriving Numerical National Water Quality Criteria for the Protection of Aquatic Organisms and Their Uses.* U.S. Environmental Protection Agency, Office of Research and Development, Duluth, MN, 98 pp.

Steevens JA, Reiss MR, and Pawlisz AV. 2005. A methodology for deriving tissue residue benchmarks for aquatic biota: A case study for fish exposed to 2,3,7,8-tetrachlorodibenzo-p-dioxin and equivalents. *Integrated Environmental Assessment and Management*. 1(2): 142-151.

Tillitt DE, Gale RW, Meadows JC, Zajicek JL, Peterman PH, Heaton SN, Jones PD, Bursian SJ, Kubiak TJ, Giesy JP, and Aulerich RJ. 1996. Dietary exposure of mink to carp from Saginaw Bay. 3. Characterization of dietary exposure to planar halogenated hydrocarbons, dioxin equivalents and biomagnification. *Environmental Science and Technology* 30:283-291.

USACE. 2004. *Environmental Residue-Effects Database*. U.S. Army Corps of Engineers, Waterways Experiment Station. Database available on line at: http://el.erdc.usace.army.mil/ered/

USEPA. 1989. *Water Quality Criteria to Protect Wildlife Resources*. Office of Research and Development. EPA/600/3-89/067. Corvallis, Oregon.

USEPA. 1995a. *Final Water Quality Guidance for the Great Lakes System; Final Rule*. Federal Register. V. 60, No. 56, pp. 15366-15425. March 23.

USEPA. 1995b. *Great Lakes Water Quality Initiative Technical Support Document for the Procedure to Determine Bioaccumulation Factors*. Office of Water. EPA-820-B-95-005.

USEPA. 1995c. *Great Lakes Water Quality Initiative Criteria Documents for the Protection of Wildlife: DDT, Mercury, 2,3,7,8-TCDD, PCBs*. Office of Water. EPA/820/B-95/008. Washington, D.C.

USEPA. 1995d. *Great Lakes Water Quality Initiative Technical Support Document for Wildlife Criteria*. Office of Water. EPA/820/B-95/009. Washington, D.C.

USEPA. 1996. *Use of Uncertainty Factors in Toxicity Extrapolations Involving Terrestrial Wildlife*. Office of Research and Development, Office of Water. Peer Reviewed Draft, Washington, DC

USEPA. 1998. *Guidelines for Ecological Risk Assessment.* U.S. Environmental Protection Agency, Risk Assessment Forum, Washington, DC, EPA/630/R095/002F. (address as of 1-31-05: http://cfpub.epa.gov/ncea/raf/recordisplay.cfm?deid=12460)

USEPA. 2000. *Methodology for Deriving Ambient Water Quality Criteria for the Protection of Human Health (2000).* Office of Water, EPA-822-B-00-004. Washington, DC. (available at: http://www.epa.gov/waterscience/criteria/humanhealth/method/index.html)

USEPA. 2002. *Trophic Level and Exposure Analyses for Selected Piscivorous Birds and Mammals: Volumes 1, 2 and 3*. Office of Water, Washington, DC.

USEPA. 2003. *Methodology for Deriving Ambient Water Quality Criteria for the Protection of Human Health (2000), Technical Support Document Volume 2: Development of National Bioaccumulation Factors*. Office of Water. EPA-822-R-03-030. Washington, D.C.

USEPA. 2004. *Draft Aquatic Life Water Quality Criteria for Selenium-2004*. Office of Water, EPA-822-D-04-001. Washington, DC. November. (available at: http://www.epa.gov/waterscience/criteria/selenium/index.htm)

USEPA Science Advisory Board. 1994. Advisory on the development of a national wildlife criteria program. USEPA Science Advisory Board, Washington, DC. EPA-SAB-EPEC-ADV-94-001.

U.S. Fish and Wildlife Service. 2003. *Evaluation of the Clean Water Act Section 304(a) human health criterion for methylmercury: protectiveness for threatened and endangered wildlife in California*. U.S. Fish and Wildlife Service, Sacramento Fish and Wildlife Office, Environmental Contaminants Division. Sacramento, California. 96 pp + appendix.

Veith GD, DeFoe, DL, Bergstedt, BV. 1979. Measuring and estimating the bioconcentration factor of chemicals in fish. *Journal of the Fisheries Research Board Canada* 36:1040-1048.

Veith GD, Call DJ, and Brooke L. 1983. Structure-toxicity relationships for the fathead minnow, Pimephales promelas: Narcotic industrial chemicals. *Canadian Journal of Fisheries and Aquatic Science* 40:743-748.

Verhaar, HJM., De Wolf, W, Dyer, S, Legierse, KCHM, Seinen, W, Hermens, JLM. 1999. An LC_{50} vs time model for aquatic toxicity of reactive and receptor-mediated compounds. Consequences for bioconcentration kinetics and risk assessment. *Environmental Science and Technology*. 33:758-763.

7 Appendix A

Cursory Review of Tissue Concentration-Response Data for Aquatic Organisms Contained in Two Databases

This appendix contains results from initial queries made of two databases containing tissue concentration-response data for aquatic organisms. These databases are:

- Jarvinen and Ankley (1999)
- Environmental Residue-Effects Database (USACE) (downloaded on September 27, 2004)

The primary purpose of these queries was simply to characterize basic attributes of the available tissue concentration-response data that have been coded to date. For example:

- How many species are represented by various chemical?
- How frequent are different endpoint classes represented? (e.g., mortatlity, growth, reproduction)
- How frequent are different types of tissues represented?
- What are the most commonly tested species?
- What exposure routes are most commonly tested?

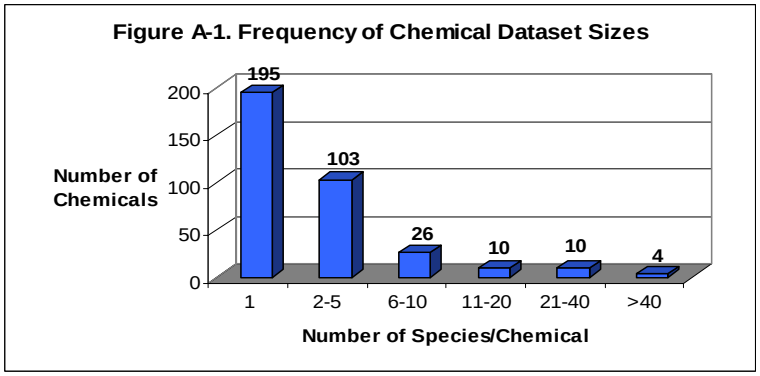
Both databases were available in electronic formats and were merged into a single MS Access™ database for further analysis. Duplicate records were removed when unambiguous determinations could be made. The ERED database contained significantly fewer fields than the Jarvinen and Ankley database, thus a number of fields in the merged database were unpopulated.

A few important caveats should be noted:

1. No attempt was made to review or screen the data for quality purposes.
2. In many cases, records reflect multiple effect levels of a given endpoint from the same test (e.g., NOAELs, LOAELs, and ECxx were recorded as separate records). The database structure did not enable unambiguous identification of "paired" NOAELs and LOAELs.
3. Nomenclature for classifying data between the two databases differed in some cases. Original classification was retained in situations were interpretations of nomenclature differed.
4. As a result of these and other factors, the actual amount of data that would be useful for criteria derivation purposes in the merged database would likely be significantly less than represented here (i.e., data were not screened). However, newer data not captured by these databases may mitigate the reduction in useable data to some extent.

Based on these simple queries, the following statements can be made regarding the status of the coded tissue concentration-response data:

1. The vast majority of chemicals are represented by 5 or fewer aquatic species (about 85%). Only about 7% of chemicals coded in the database contain more than 10 aquatic species represented (Figure A-1).



2. Organic chemicals with 6 or more species represented include:

Organochlorine Pesticides:

aldrin, DDE, DDT, chlordane, endrin, endosulfan, heptachlor, kepone, lindane, methoxychlor, mirex, toxaphene

Organophosphate Pesticides:

chlorpyrifos, diazinon

Pyrethroids:

fenvalerate, permethrin

PAHs:

benzo(a)pyrene, flouranthene, naphthalene, phenanthrene, pyrene

Metals/metalloids:

arsenic, cadmium, chromium, copper, lead, mercury, nickel, selenium, silver, vanadium, zinc

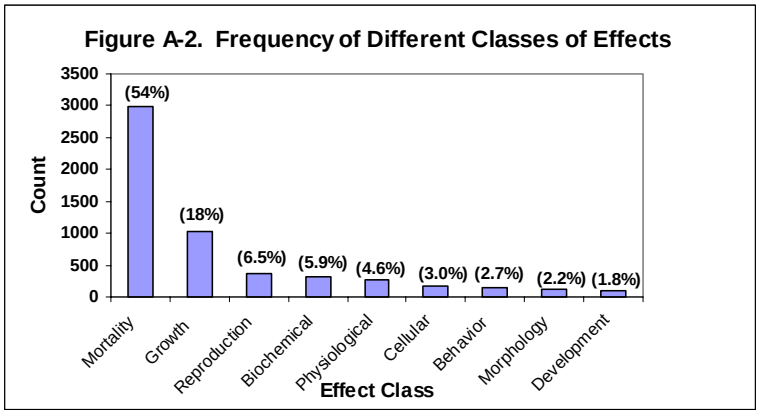
PCBs/Dioxins:

2,3,7,8-TCDD, various aroclors

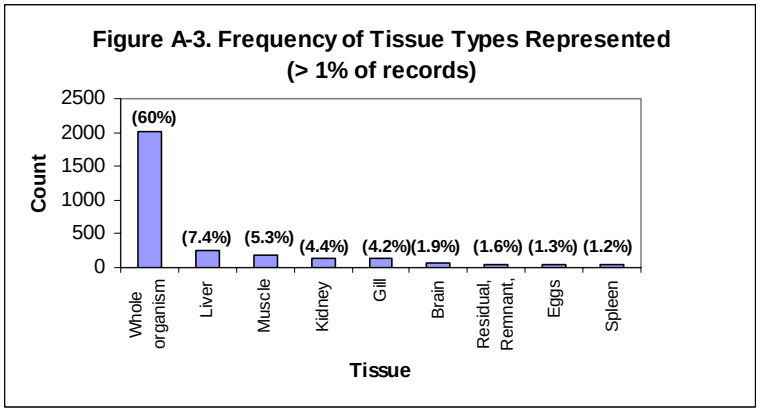
Other:

PCP, 2,4,6-trichlorophenol, 2,3,4,6-trichlorophenol, hexachlorobenzene, pentachlrobenzene, TBT, dibutyltin, di-2-ethylhexyl phthalate

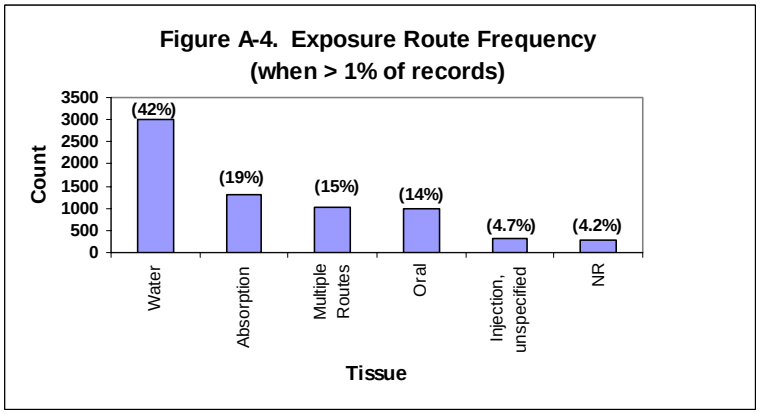
3. Mortality is by far the most common endpoint measured (over half of the coded data). Reproductive endpoints constitute a relatively small fraction of the data (about 6%; Figure A-2)



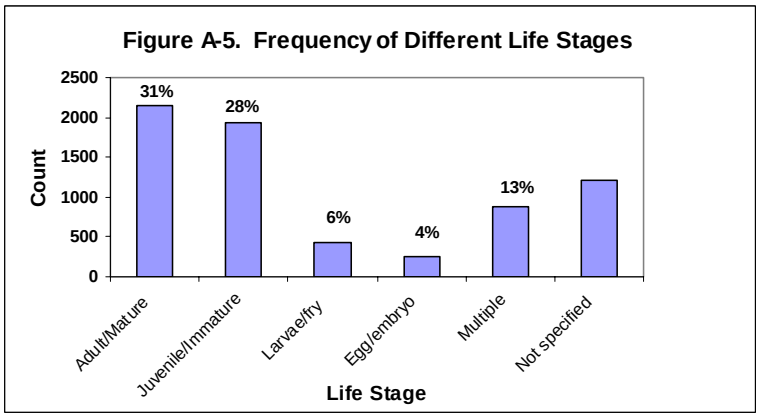
4. Whole organism measurements are by far the most common tissue sampled (Figure A-3).



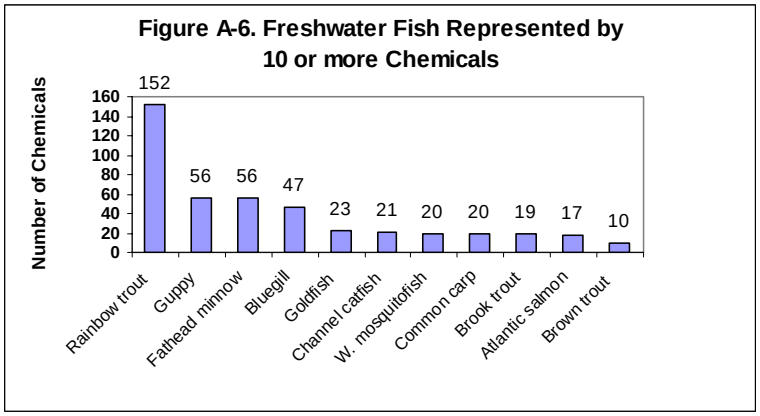
5. Water only exposures are most common, followed by multiple routes and oral (presumably food ingestion; Figure A-4).



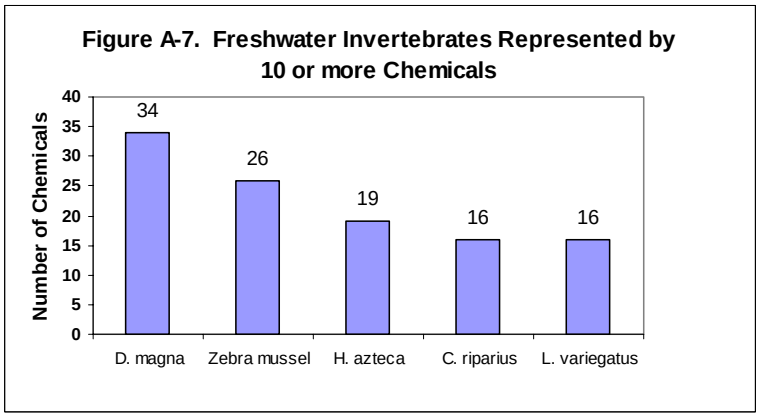
6. Adult and juvenile/immature life stages are most commonly represented (about 60% of the records), with early life stages (larval/fry and egg/embryo) and multiple life stages constituting about 10% and 13% of the records, respectively (Figure A-5).



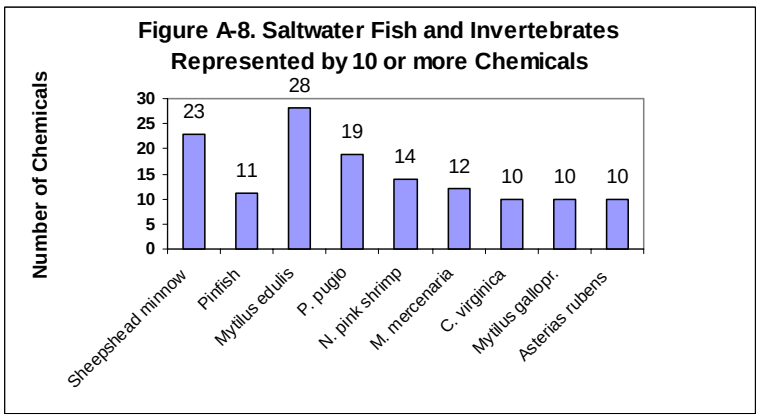
7. Most common groups of freshwater fish species represented are salmonids (rainbow trout, brook trout, brown trout, atlantic salmon), followed by cyprinids (fathead minnow, goldfish, common carp) and Poeciliidae (guppy, mosquitofish; Figure A-6).



8. The most common freshwater invertebrates represented include a cladoceran (*D. magna*), an amphipod (*H. azteca*), a mollusk (zebra mussel), an insect (*C. riparius*) and an oligochaete (*L. variegates*; Figure A-7).



9. Few saltwater fish species are represented broadly in the database (i.e., for 10 or more chemicals) while shrimp and bivalve mollusks are among most commonly tested invertebrate species (Figure A-8).



References for Appendix A

Jarvinen, AW, Ankley, JT. 1999. *Linkage of Effects to Tissue Residues: Development of a Comprehensive Database for Aquatic Organisms Exposed to Inorganic and Organic Chemicals*. SETAC Press, Pensacola, FL, 364 pp.

USACE. 2004. Environmental Residue-Effects Database. U.S. Army Corps of Engineers, Waterways Experiment Station. Database available on line at: http://el.erdc.usace.army.mil/ered/