♻EPA AN SAB REPORT: EVALUATION OF THE GUIDANCE FOR THE GREAT LAKES WATER QUALITY INITIATIVE

PREPARED JOINTLY BY THE GREAT LAKES WATER QUALITY SUBCOMMITTEE OF THE ECOLOGICAL PROCESSES AND EFFECTS COMMITTEE AND THE DRINKING WATER COMMITTEE

December 16, 1992

EPA-SAB-EPEC/DWC-93-005

Honorable William Reilly
Administrator
U.S. Environmental Protection Agency
401 M Street, S.W.
Washington, D.C. 20460

Subject: SAB Review of Guidance for the Great Lakes Water Quality Initiative

Dear Mr. Reilly:

The Science Advisory Board (SAB) has completed its review of four technical guidance documents for developing water quality criteria in the Great Lakes Basin. This guidance was developed by EPA in collaboration with states in the Great Lakes Basin and is intended for application in this region. This review covered a wide range of disciplines and included the expertise of both the Ecological Processes and Effects Committee (EPEC) and the Drinking Water Committee (DWC) of the SAB. The SAB conducted this review in response to an extensive charge from EPA Region V which asked for review on aquatic life, wildlife, and human health criteria guidance and a new approach for assessing bioaccumulation. EPEC formed a Great Lakes Water Quality Subcommittee to coordinate the review. The Subcommittee evaluation focused on the following issues from the charge: 1) the validity and proposed uses of Tier 2 aquatic life criteria; 2) the wildlife criteria approach, including species selected, the data used, and the use of Toxicity Equivalency Factors (TEFs); and 3) the use and calculation of bioaccumulation factors. The Subcommittee and the Drinking Water Committee addressed the human health criteria for carcinogens and minimum data sets for each tier.

Four public meetings were conducted, including two meetings by the Drinking Water Committee (focused on Human Health Criteria) and a meeting of the Dioxin Ecotox Subcommittee of EPEC which included TEFs for aquatic life and wildlife.

The Great Lakes Water Quality Initiative (GLWQI) is a challenging and ambitious endeavor. The SAB commends EPA for the interactions among the states, EPA, the private sector and the scientific community in further developing environmental protection programs for the Great Lakes. This program should also actively involve interests in Canada and seek a consistent U.S.-Canadian approach. Based on the documents reviewed and the presentations made to the panels it is unclear how the Great Lakes region is unique in water quality problems and issues. The Subcommittee recommends that the EPA provide more background information on the sources and effects of chemicals discharged to the Great Lakes and the nature of the exposures. The GLWQI should revise its introduction to discuss its rationale for an initiative in the Great Lakes; a history of contaminant related ecological problems; and a discussion of environmental issues associated with the Great Lakes. These would be of value to place the GLWQI in perspective.

The Subcommittee also recommends that the EPA promote a broadly based ecosystem approach which considers not only point source discharges but non-point sources, sediments, atmospheric fall-out, and groundwater as targets for conservation and control of undesirable loadings (i.e., levels which have a toxic effect). Likewise, the EPA should consider other pathways of exposure and endpoints of effects for wildlife and humans.

The Subcommittee supports the principle of using Tier 1 and Tier 2 data in developing water quality protective of aquatic life, wildlife and humans. The Tier 1 criteria have data sets equivalent to the National Water Quality Criteria. The Tier 2 approach is used to develop criteria for contaminants which have less data. The Subcommittee recommends that the Tier 2 minimum data base for aquatic life include estimates of chronic toxicity and assess matrix effects on toxicity. We caution EPA against setting inflexible numeric standards based on Tier 2. Tier 2 derived values should be used as an incentive to improve the underlying data base for Tier 2 chemicals.

The Subcommittee supports the GLWQI's efforts to develop an approach to protect wildlife from the effects of bioaccumulative chemicals in the environment. However, the Subcommittee is concerned that the current approach does not adequately consider ecologically important species in selection of surrogate wildlife species and it relies on human health procedures that are more appropriate for protection of individuals than for local or regional wildlife populations. Similarly, the Subcommittee feels that the definition of wildlife is ambiguous and

2

recommends that EPA and the GLWQI develop a definition of wildlife and justify species inclusions and exclusions.

The form of the contaminant and the analytical methods to measure criteria concentrations deserves further discussion in the guidance. The Subcommittee recommends that values for both the biologically active form of contaminants and the total concentration be included in water quality criteria. Guidance should be provided for monitoring instances where water quality standards result in water concentrations that are well below detection limits of currently accepted analytical methods.

The Subcommittee notes that the GLWQI appears to have no elements which predict the persistence of chemicals. The proposed approaches also do not consider rates of degradation, hydrolysis, volatilization, sorption and all of the environmental transport and fate pathways. The approach for assessing bioaccumulation factors (BAF) advanced in the GLWQI uses octanol/water partition coefficients (Log P) and food chain models to predict residues in biota. Other approaches to estimate persistence should also be explored, such as, using biological residues and partitioning methods with C_{18} and/or Tenax.

We are concerned that the Great Lakes Initiative human health risk assessment methodology is not using the most updated approaches being used by EPA and others. Tier 1 criteria for human health should be limited to chemicals with good data on carcinogenesis, reproductive and developmental/ teratogenic effects. The linear multistage model is a reasonable default methodology for chemicals which lack more detailed information on their modes of action. Ideally, additivity should not be used as a default, but rather multiple carcinogens should be considered on a case by case basis. We encourage EPA to use a variety of broad criteria to classify chemicals as Tier 2 to encourage improvements in the data base. The Subcommittee recommends that the draft human health criteria documents and guidance for their development be revised to improve the analysis and presentation of data and rationale for the development of the criteria.

It is the SAB's understanding that the draft guidance and implementation procedures will be published in the Federal Register for public comment. It is the Subcommittee's conclusion that the substantive scientific issues raised here should be addressed before the Agency adopts final guidance. The SAB would like the opportunity to review the revised guidance and public comments prior to the final publication. We are particularly interested that the Agency respond to our

3

recommendations for expanding the data set for Tier 2 aquatic life criteria, the population approach for wildlife, the data requirements for human health Tier 1 and handling multiple carcinogens, the relationship of the GLWQI to other media within an ecosystematic context, and the Agency's plans for implementation of the guidance. We appreciate the opportunity to review this important Agency initiative and look forward to receiving your response.

Sincerely yours,

Dr. Raymond C. Loehr, Chair
Executive Committee
Science Advisory Board

Dr. Kenneth L. Dickson, Chair
Great Lakes Water Quality Subcommittee
and Ecological Processes and
Effects Committee
Science Advisory Board

Dr. Verne Ray, Chair
Drinking Water Committee
Science Advisory Board

NOTICE

This report has been written as a part of the activities of the Science Advisory Board, a public advisory group providing extramural scientific information and advice to the Administrator and other officials of the Environmental Protection Agency. The Board is structured to provide a balanced expert assessment of scientific matters related to problems facing the Agency. This report has not been reviewed for approval by the Agency; and hence, the contents of this report do not necessarily represent the views and policies of the Environmental Protection Agency or other agencies in Federal government. Mention of trade names or commercial products does not constitute a recommendation for use.

i

ABSTRACT

The report represents the conclusions and recommendations of the U.S. Environmental Protection Agency's Science Advisory Board (SAB) regarding a EPA guidance for the Great Lakes Water Quality Initiative (GLWQI). The SAB commends the Agency for the interactions among the states, EPA, the private sector and the scientific community which have lead to the development of this initiative. The SAB recommended that the introduction to the guidance be revised to explain the unique characteristics of the Great Lakes and the rationale for an initiative. The SAB endorsed the ecosystems approach of the initiative and recommended that it also address non-point sources, atmospheric deposition and contaminated sediments. The Subcommittee agreed with the concept of Tier 1 and Tier 2 criteria but was concerned that the minimal data base currently required in Tier 2 water quality criterion - a single acute toxicity test - is inadequate. They were also concerned that the risk management apparatus currently in place; cf., the anti-backsliding provisions of the Clean Water Act, may prevent adjustments in Tier 2 numbers when more data become available. The Subcommittee recommends that the approach to protect wildlife be expanded to consider ecologically representative species and species sensitivities and to focus on populations. The current wildlife criteria concepts were formulated around the perceived requirements of the human health risk assessment paradigm and they are inadequate for wildlife. The Subcommittee recommended that the program also consider both the biologically active form and the total contaminants concentrations when establishing water quality criteria. The GLWQI should provide some specific guidance on how to handle monitoring compliance for criteria which are below the detection limits of analytical methods. The Subcommittee recommended that the GLWQI add procedures to predict the persistence of chemicals.

The SAB is concerned that the human health risk assessment methodology being advanced by the GLWQI is not using updated approaches for exposure assessment and carcinogen classification that are being used by EPA and others. Tier 1 should be limited to chemicals with good data on carcinogenesis, reproductive and developmental/teratogenic effects. The linear multistage model is a reasonable default methodology for chemicals which lack more detailed information on their modes of action. Ideally, multiple carcinogens should be considered on a case by case basis. The SAB encouraged EPA to use a variety of broad criteria to classify chemicals as Tier 2 to encourage improvements in the data base. The SAB recommended that the draft human health criteria documents

and guidance for their development be revised to reflect SAB comments and improve the analysis and presentation of data and rationale for the development of the criteria.

U.S. ENVIRONMENTAL PROTECTION AGENCY
SCIENCE ADVISORY BOARD
GREAT LAKES WATER QUALITY SUBCOMMITTEE

SCIENCE ADVISORY BOARD STAFF

Dr. Edward S. Bender, Biologist and Designated Federal Official, Science Advisory Board, 401 M Street, S.W., A-101F, Washington, D.C.

Mrs. Marcia K. Jolly, Secretary to the Designated Federal Official

ENVIRONMENTAL PROTECTION AGENCY
SCIENCE ADVISORY BOARD
DRINKING WATER COMMITTEE

CHAIR

Dr. Verne A. Ray, Medical Research Laboratory, Pfizer Inc., Groton, Connecticut

VICE-CHAIR

Dr. Vern L. Snoeyink, Department of Civil Engineering, University of Illinois, Urbana, Illinois

MEMBERS/CONSULTANTS

Dr. Richard J. Bull, College of Pharmacy, Washington State University, Pullman, Washington

Dr. Gary P. Carlson, Department of Pharmacology and Toxicology, School of Pharmacy, Purdue University, West Lafayette, Indiana

Dr. Keith E. Carns, East Bay Municipal Utility District, Oakland, California

Dr. Lenore S. Clesceri, Rensselaer Polytechnic Institute, Troy, New York

Dr. David G. Kaufman, Department of Pathology, University of North Carolina, Chapel Hill, North Carolina

Dr. Ramon G. Lee, American Water Works Service Company, Voorhees, New Jersey

Dr. Edo D. Pellizzari, Research Triangle Institute, Research Triangle Park, North Carolina

Dr. Mark D. Sobsey, Department of Environmental Sciences and Engineering, School of Public Health, University of North Carolina, Chapel Hill, North Carolina

Dr. James M. Symons, Department of Civil and Environmental Engineering, University of Houston, Houston, Texas

SCIENCE ADVISORY BOARD STAFF

Mr. A. Robert Flaak, Assistant Staff Director and Acting Designated Federal Official, Science Advisory Board (A-101F), U.S. EPA, 401 M Street, SW, Washington, DC 20460

Mrs. Frances Dolby, Staff Secretary, Drinking Water Committee, Science Advisory Board (A-101F), U.S. EPA, 401 M Street, SW, Washington, DC 20460

TABLE OF CONTENTS

1. EXECUTIVE SUMMARY

The Great Lakes Water Quality Subcommittee of the SAB Ecological Processes and Effects Committee (EPEC) was asked to review the scientific underpinnings of the proposed Great Lakes Water Quality Initiative (GLWQI). The Subcommittee met February 18-20, 1992 for briefings on the technical approaches for developing water quality criteria for aquatic life, wildlife, and human health in the Great Lakes and to receive public comments. This report summarizes the findings of the Subcommittee which addressed all parts of the guidance and the recommendations of the Drinking Water Committee which focused on the Human Health Criteria (see Chapter 6 of this report).

The Great Lakes Water Quality Initiative (GLWQI) is a challenging and ambitious endeavor. In addition to parties in the United States, Canadian interests must also become actively involved in developing a consistent approach for this shared international resource. Based on the materials reviewed and the presentations to both Panels, it is evident that a great deal of time and effort has been spent by all parties. The SAB encourages the continuation of interactions among the states, EPA, the private sector and the scientific community in further developing environmental protection programs for the Great Lakes.

The SAB recommends that the introduction to the documents be revised to explain how the Great Lakes are unique in terms of their water quality problems and issues, and indicate how the unique aspects of contaminant exposure of the biota in the Great Lakes dictate the approach being advanced. A better rationale should be developed and presented in the guidance documents on why a Great Lakes specific approach for establishing water quality criteria for aquatic life, wildlife and humans is needed. Inclusion of data showing trends in the levels of contaminants in the Great Lakes and a history of contaminant related ecological problems and issues associated with the Great Lakes would be of value to place the proposed program in perspective.

The GLWQI makes an effort to use an ecosystem approach to environmental protection. The Subcommittee strongly endorses an ecosystems approach because it is more scientifically sound than the piece-meal approach that has been historically used. The approach should also take into account the sources, sinks, and transport routes of these chemicals. The great opportunity of an ecosystem approach is to capture the major inputs and target resources for the most effective

1

control measures. It is not clear however, what specific mechanisms the GLWQI has incorporated to address non-point sources, atmospheric deposition and contaminated sediments. The current approach is specifically directed at point sources effects on water quality and biota. A complete ecosystem approach should examine all sources of contaminant loadings, all ecosystem compartments and all ecological receptors. While the SAB recognizes that this is difficult, the position of the GLWQI in an overall ecosystem approach for environmental protection of the Great Lakes should be identified.

A fundamental aspect of the Great Lakes Water Quality Initiative is the principle of using Tier 1 and Tier 2 data in developing water quality protective of aquatic life, wildlife and humans. The Subcommittee agrees in concept with this approach. There are many chemicals for which Tier 1 data do not exist, yet which need regulation. A Tier 2 approach, if properly applied, provides a mechanism for controlling those chemicals for which there are limited scientific data while at the same time provides a mechanism for reducing uncertainty regarding their environmental consequences. The Subcommittee is concerned that the minimal data base currently required in Tier 2 - a single acute toxicity test - is inadequate. The Subcommittee recommended that the Tier 2 minimum data base include estimates of chronic toxicity and matrix effects in toxicity.

The SAB fully expects that as additional scientific data accumulate, the technically derived values for WQC will change in response to new information. It is not clear, however, that the risk management apparatus currently in place is capable of accommodating these scientific improvements; cf., the anti-backsliding provisions of the Clean Water Act. The Board is concerned that situations could arise in which risk management positions are not scientifically defensible.

The Subcommittee supports the GLWQI's efforts to develop an approach to protect wildlife from the effects of bioaccumulative chemicals in the environment. However, the Subcommittee is concerned that the current approach does not adequately consider ecologically representative species in selection of surrogate wildlife species. Similarly, the Subcommittee feels that the definition of wildlife is ambiguous as used in the GLWQI. We recommend that EPA and the GLWQI develop a definition of wildlife and justify species inclusions and exclusions. Regardless of the definition, provisions should be provided in the GLWQI for re-evaluating and updating the list of surrogate species. In addition, the exposure assessment needs to be differentiated between species sensitivities and effects of the chemicals.

2

The Subcommittee is also concerned that the methodology used in the GLWQI to assess the range of species sensitivities needs further development. In contrast to human health criteria which are designed to protect individuals, wildlife criteria are designed to protect populations and must consider differences in species sensitivities. This aspect is not a part of the human health methodology which has been applied to establish wildlife criteria in the GLWQI. The discussions of the Lowest Observed Acute Effect Level (LOAEL) versus the No Observed Acute Effect Level (NOAEL) in the Technical Support Document are very superficial. These concepts were formulated around the perceived requirements of the human health risk assessment paradigm. While they may be applicable to human health risk assessment, they cannot serve as foundations for the development of criteria methodologies for the protection of wildlife. Further explanation is needed of how the two applications differ and how they will be addressed. The GLWQI should develop guidance for the selection of NOAELs appropriate for the protection of local and regional wildlife populations as distinct from the protection of individuals.

There are a number of regulatory approaches which have direct bearing on chemical, physical and biological water quality and protection of aquatic life, wildlife and humans. The relationship of these approaches to the GLWQI is unclear and was not adequately addressed in the materials examined by the Subcommittee. What is the relationship of the proposed GLWQI to whole effluent biomonitoring? How does the GLWQI focus on bioaccumulative chemicals relate to the HPLC based screening approaches for bioaccumulative chemicals in effluents? What is the interface between the GLWQI and the National Sediment Quality program? It is not clear from the documents reviewed and the presentations that these techniques and activities were considered in developing the GLWQI approaches.

The GLWQI is designed to establish water quality criteria for total contaminant concentration not the bioavailable form of the contaminant. The Subcommittee recommends that the program also consider the biologically active form of contaminants when establishing water quality criteria. The Subcommittee feels that by basing the water quality criteria only on total concentration that much of the science which has developed in the last ten years on the importance of chemical speciation and biological activity is being ignored. The approach is also inconsistent with the one for sediment criteria which uses the soluble forms of contaminants, but not the total concentration.

Since a water quality criterion for a chemical may be the result of a back calculation from a measured or predicted biological residue concentration and/or based on Tier 2 data, the appropriately calculated water concentration may be several orders of magnitude below detection limits of currently accepted analytical methods. This creates a serious compliance monitoring problem and may further widen the credibility gap between the regulatory agencies, regulated community and the public. The GLWQI should provide some specific guidance on how to handle this problem.

The Subcommittee notes that the GLWQI appears to have no elements which predict the persistence of chemicals. The proposed approaches also do not consider rates of oxidation, hydrolysis, volatilization, sorption and all of the environmental transport and fate pathways. The approach for assessing bioaccumulation factors (BAF) advanced in the GLWQI uses octanol/water partition coefficients (Log P) and food chain models to predict residues in biota. While these approaches appear to have some utility, there are alternative approaches (Lebo et al., 1992 and Johnson, 1991) which should also be explored as part of the GLWQI such as using biological residues, partitioning methods with C_{18} and/or Tenax and "artificial fish".

The SAB is concerned that the human health risk assessment methodology being advanced by the GLWQI is not using updated approaches for exposure assessment and carcinogen classification that are being used by EPA and others. Tier 1 should be limited to chemicals with good data on carcinogenesis, reproductive and developmental/teratogenic effects. Tier 2 should contain chemicals for which a less complete data set exists and appropriate uncertainty factors are incorporated to compensate for this lack of data. The Agency must move forward by using biologically based models for assessing carcinogenic risks at low doses. The linear multistage model is a reasonable default methodology, but the Agency appears reluctant to follow its own guidelines when appropriate mechanistic, pharmacokinetic, or other relevant data are available for individual chemicals. Ideally multiple carcinogens should be considered on a case by case basis, because the assumption of additivity has both practical and scientific shortcomings. The SAB recommends that the draft human health criteria documents and guidance for their development be revised to reflect SAB comments and improve the analysis and presentation of data and rationale for the development of the criteria.

4

2. INTRODUCTION

The Great Lakes Water Quality Initiative (GLWQI) was developed by the U.S. EPA in cooperation with the states in the Great Lakes Basin to ensure consistency in the development of water quality standards to maintain, protect, and restore the unique Great Lakes resource. EPA Region 5 (Chicago, IL) has taken the lead role since the effort began in 1989. The schedule for GLWQI activities, under the Great Lakes Critical Programs Act, required EPA to publish final GLWQI guidance by June 1992 and for Great Lakes States to adopt the guidance as part of their state water quality standards within two years of the publication of final guidance.

The GLWQI consists of six interconnected procedures: a) derivation of criteria for the protection of aquatic life; b) bioaccumulation factors; c) derivation of criteria for the protection of wildlife; d) derivation of criteria for the protection of human health; e) protection of current water quality (antidegradation); and f) translation of standards into regulatory controls (implementation). This guidance was developed by technical work groups of scientists from the states, U.S. EPA, U.S. Fish and Wildlife Service, and the U.S. National Park Service with input from a public participation group which includes members of the regulated community and academia. As a result the GLWQI is being developed and implemented through an iterative process and has a goal of being based on a broad consensus.

The GLWQI guidance is being developed primarily from the water quality perspective and it will be implemented through the National Pollutant Discharge Elimination System (NPDES) permit program. As a result, it draws heavily from the Agency's technical guidance and experience with water quality criteria and surface water monitoring. States rely on the water quality criteria as the foundation of their state water quality standards. The GLWQI is an effort to coordinate the surface water regulatory needs within a geographically-similar area. In theory, such initiatives could address regulatory needs for a variety of media and emphasize the special environmental problems of the region. This GLWQI guidance includes a process to set environmental quality criteria using a smaller data set (Tier 2) than national guidelines. The guidance also introduces guidelines for developing wildlife criteria and a revised process for estimating the potential of chemicals to bioaccumulate.

2.1   Statement of the Charge

On January 8, 1992, Mr. Dale Bryson, Director, Water Division, EPA Region 5, sent a revised charge to the Science Advisory Board (SAB) requesting a review of the recommendations of the EPA-State Great Lakes Water Quality Initiative. In particular, he asked that the SAB focus on questions related to four guidance documents: Aquatic Life; Bioaccumulation Factors; Wildlife Methodology; and Human Health.

Specific aspects of the charge were:

A.   Tier 2 Aquatic Life Proposal

1)   Does the Tier 2 methodology provide a valid method for developing values in the absence of sufficient data to meet the Tier 1 requirements?

2)   Is the derivation of the values compatible with the proposed uses of the Tier 2 values?

B.   Wildlife Criteria Methodology

1)   Is the wildlife criteria algorithm, which only considers dietary and drinking water exposures reasonable?

2)   Are the representative avian and mammalian species reasonable and appropriate selections?

3)   Is the general approach for using uncertainty acceptable? Is the derivation of those factors adequately explained?

4)   With regard to the four wildlife criteria calculated for Mercury, DDT, Dioxin, and PCBs, are the toxicity data reviewed complete and their subsequent interpretation appropriate?

5)   Are the TEFs [toxicity equivalent factors] chosen for dioxins, coplanar and monortho coplanar PCBs acceptable and is their application in deriving criteria adequately presented?

6

C. Bioaccumulation Factors (BAF)

1) Is the BAF/BCF [Bioconcentration Factors] for organic chemicals usefully related to the percent lipid in tissues?

2) Are field measured BAFs suitable for the calculation of generally applicable criteria?

3) Is a BCF an underestimate of a BAF for organic chemicals with log K_{ow} in the range of 4.5 to 6.5?

4) Is the proposal to adjust BCFs to BAFs, based on Thomann (1989) appropriate?

5) Are there chemicals or groups of chemicals (e.g., PAHs?) with log K_{ow} in the 4.5-6.5 range for which the application of a food chain multiplier is not appropriate?

D. Human Health Criteria

1) Is the Linearized Multistage Model appropriate to manage chemical carcinogens?

2) Should additive risk be considered in evaluating ambient water quality of the Great Lakes? If so, how should this be presented?

3) Is the proposed minimum data set for Tier 1 appropriate for establishing region wide numeric water quality criteria? Is it appropriate to treat A and B level carcinogens and certain designated C level carcinogens equally via this approach?

4) Is it defensible to regulate environmental contaminants via the proposed Tier 2 approach? If so, what is the minimum database necessary? Is it scientifically defensible to manage C level carcinogens via this concept?

5). Is the approach to identify chemical mutagens and teratogens for special consideration appropriate? If so, how should they be controlled?

6) Is the concept of relative source contribution a scientifically valid approach in controlling the significance of this route in total food chain and other exposure to bioaccumulative contaminants?

The Subcommittee accepted this charge, but also requested further information on the use and rationale for these procedures and clarification of certain regulatory definitions that would affect these concepts. The Subcommittee agreed that the human health criteria guidance would also be reviewed by the Drinking Water Committee of the SAB and the application of TEFs would be reviewed by the SAB's Dioxin Ecotox Subcommittee.

2.2 Subcommittee Review Procedures

The Ecological Processes and Effects Committee (EPEC) of the SAB was assigned the lead for coordinating the review of the GLWQI technical guidance. The Great Lakes Water Quality Subcommittee was composed of members and consultants from EPEC with expertise to address the four areas of the charge from the perspective of surface water quality. Two members of the Drinking Water Committee provided additional expertise on human health criteria and human cancer risk assessment methodologies to the Subcommittee. In addition, the Drinking Water Committee, including a consultant from EPEC, separately reviewed the Human Health Criteria guidance.

The Great Lakes Water Quality Subcommittee (GLWQS) met in Rosemont, Illinois on February 18-20, 1992 to receive briefings on the technical guidance and take public comments. The Chairman summarized the preliminary impressions of the Subcommittee on the overall initiative. He also explained that other panels of the SAB would address portions of the charge. At the meeting, the Chairman asked EPA to provide the Subcommittee with further information regarding implementation of the guidance (January 1992 version of the Federal Register preamble to the guidance), the goals of the program and its relationships to other media, the reasons for a unique approach in the Great Lakes Basin, and the process for monitoring compliance of criteria that are below analytical levels of detection. The GLWQI provided the Subcommittee with copies of the draft

preamble (dated January, 1992), implementation guidance (created December, 1991), and other documents to address these issues. In addition, the Subcommittee received written comments from 16 parties and heard oral comments at the meeting. In May 1992, the Subcommittee held a writing session.

The Drinking Water Committee met on April 14 and June 1, 1992 to review the guidance for human health criteria. Copies of comments from the GLWQS were provided to that group and three members of the GLWQ Subcommittee were present for the discussion. The comments of the DWC appear primarily in Chapter 6 of this report.

The Subcommittee also received input from the SAB's Dioxin Ecotox Subcommittee of EPEC which reviewed a question related to Toxicity Equivalent Factors (TEFs). The Subcommittee noted that the use of TEFs would be addressed by the SAB's Environmental Health Committee (EHC) as part of the reevaluation of the Agency's Dioxin Risk Assessment, therefore, the comments here and in the Dioxin Ecotox review (Science Advisory Board, 1992) were limited to specific research needs on TEFs for wildlife and aquatic life.

3. AQUATIC LIFE CRITERIA

3.1 Summary of the Proposed Tier 2 Method

A two tiered procedure to derive aquatic life water quality criteria is being proposed in the Great Lakes Water Quality Initiative for the protection of aquatic life from exposure to individual chemicals contained in point source effluent discharges to the Great Lakes.

Tier 1 acute and chronic numeric criteria will be derived using a modification of the current U. S. Environmental Protection Agency's "Guidelines for Deriving Numerical National Water Quality Criteria for the Protection of Aquatic Organisms and Their Uses." Great Lakes Basin States are expected to adopt these as numeric criteria. Major modifications of the original EPA guidelines included the following: deletion of saltwater criteria, final residue value, and considerations related to wildlife species; and inclusion of lower criteria to protect commercially or recreationally important species in the Great Lakes Basin, use of Ceriodaphnia 7-day life cycle test in the criteria, and the use of the two tiered approach when sufficiently large data bases are not available.

The Tier 2 approach is structured in a manner conceptually similar to the U.S. EPA water quality criteria method. A statistical procedure was applied to the "universe" of data existing within the EPA water quality data base. Instead of the Tier 1 requirement of a minimum base of acute toxicity test results from 8 species, according to the proposal, Tier 2 can be used if the acute toxicity data base includes results from a single species of daphnid. Daphnids were included in the data base because they appeared to be the most sensitive species for many persistent chemicals. This use of limited data requires that an uncertainty factor be used to account for the variability associated with toxicological responses, laboratory testing methods, and extrapolations to the real world. A gradient of species variability uncertainty factors and acute to chronic ratio uncertainty factors emerged from the relationships defined in the statistical analyses of the complete Tier 1 water quality criterion data sets. Selection of the uncertainty factors from the 50th to 99th percentile is a policy issue. How the Tier 2 data are used is critical to the entire concept of "short-cut-methods" to derive criteria or values.

The GLWQI proposed Tier 2 approach was developed to permit development of criteria and standards for chemicals, when the aquatic toxicity data base was

10

not adequate to meet the Tier 1 data base requirements. GLWQI anticipates the Tier 2 criteria concentrations will be more stringent, i.e. over-protective, than the Tier 1 criteria. Host et al. (1990), developers of the statistical approach advanced for Tier 2, intended the Tier 2 values to be used as narrative standards and not to be used for numerical criteria.

The existing data base for Tier 1 chemicals was subjected to a statistical analysis to determine the effect upon the calculated Final Acute Value (FAV), if portions of the toxicity data was sequentially removed. This data set was then subjected to a probability analysis to determine if a Secondary Acute Value (SAV) was calculated according to Tier 2 protocol if there were a sequential reduction in the number of acute toxicity values available, i.e., from 8 to 7, 7 to 6, ..., 2 to 1. The resulting data sets were then analyzed to determine the percent frequency when the Tier 2 SAV would be less than the Tier 1 FAV. The choice of the percent probability used to choose the secondary acute value, i.e., 80% was based upon the assumption that the probability of the Tier 2 SAV would exceed the FAV only 20% of the time. (The Subcommittee is not aware of any rationale provided in the documentation for this particular value.) The ratio of FAV/SAV was calculated to be 3.6, which was labeled the "Final Acute Value Factor" (FAVF). This FAVF is utilized to calculate a SAV by dividing the lowest genus (Daphnid species) mean acute value (GMAV) by the FAVF; SAV = (lowest GMAV)/FAVF. The final standard, "Secondary Chronic Value" (SCV) is then calculated using the SAV divided by the Secondary Acute Chronic Ratio (SACR), which is derived from the ordered ratio's of the Tier 1 data set FAV/FACR. The SCVs were ordered from high to low so that a secondary acute-chronic ratio could be derived to correspond to any selected percentile.

3.2 Specific Responses to the Charge

3.2.1 Validity of Tier 2 Values

The intent of developing a Tier 2 protocol was to supplement the acute toxicity data base, and produce values to be adopted as narrative standards. The Tier 2 numbers were designed to provide a:

a)   basis for evaluating potential for concern,
b)   focus on chemicals which need more toxicity data,
c)   basis for regulatory limits under some circumstances.

11

The Subcommittee endorses the original intent of using Tier 2 numbers to identify those contaminants of concern which need additional toxicity data. However, the Subcommittee is concerned that Tier 2 values might be adopted as regulatory limits for point source dischargers. The Tier 2 numbers were designed to be over protective in the arbitrary choices of percentage distributions from the original data set. These numbers should only be used as interim narrative standards not as numeric limits. Otherwise, EPA may be forced to revise it's policy on anti-backsliding (see footnote 1 on page 14 and Section 3.3.4).

Under the best of circumstances water quality criteria developed using the national guideline approach are generated from data which contain significant uncertainties. For example, the statistical variances associated with the generation of EC and LC50's are not included in the derivation. The procedures used in developing the Great Lakes Initiative (GLI) aquatic criteria are based in large part on the national criteria, i.e., an assumption has to have been made that the national criteria are correct and therefore can be modified for use with significantly smaller data sets. This may in fact be true, although it is probably also true that the smaller the data set the greater the uncertainty. EPA recognizes this and the acute factors which have been generated reflect this (Table I).

Table I

Relationship Between Data Requirements and Uncertainty

Number of Minimum Data Requirements	Acute Factor
1	20
2	13
3	8.6
4	6.5
5	5.0
6	4.0
7	3.6

There are at least two features that are disconcerting about this approach. The example which follows will be used to describe one of these.

The data on which the national water quality criterion for copper was developed includes numerous tests for which the test species was *S. magna.* The LC/EC50 values for this species range from 10 ug/L to 200 ug/L. If the procedures for the GLI were followed and the only data that were available for copper were the EC50 of 200 ug/L, then the Secondary Acute Value would be:

$$\text{Secondary Acute Value} = \frac{\text{lowest GMAV}}{\text{Acute Factor}}$$

$$= \frac{200 \text{ ug/L}}{20} = 10 \text{ ug/L}$$

If on the other hand the only data available for copper were the EC50 value of 10 ug/L, then the Secondary Acute Value would be:

$$\text{Secondary Acute Value} = \frac{\text{lowest GMAV}}{\text{Acute Factor}}$$

$$= \frac{10 \text{ ug/L}}{20} = 0.5 \text{ ug/L}$$

These data suggest that for those chemicals for which there is a significant matrix effect, significant differences in the secondary maximum concentrations and also the secondary continuous concentration can exist if only a single GMAV is available for evaluation. An alternative approach might be to dictate not only the species to be tested but also the matrix, although the number of matrix factors altering bioavailability can be extensive.

The second factor of concern about this approach is the relationship between data generation and cost. It has been suggested that the costs of generating a complete data set for deriving a National Water Quality Criteria could be a much as $100,000 per chemical. However, there has to be a gradient for costs between generating a single acute value and complete data set. If short term chronic test results are acceptable for input into the derivation of the Great Lakes criteria then tests could be undertaken for less than $5,000 per chemical for two matrices and two species of test organisms.

13

The Subcommittee is concerned that the minimal data base of one species acute test is inadequate. From a statistical perspective, the historical data base is probably scientifically defensible to account for many of the sources of toxicological testing uncertainty. However, a purely statistical analysis of the existing historical water quality data base does not reflect several important contemporary considerations. Although acute toxicity data can be very useful when there is a void of other data, the current state-of-the-science is to rely upon data that are more characteristic of chronic effects. Some new fairly inexpensive short cut methods with some plants, invertebrates, and fishes offers many advantages over acute data with extrapolations to chronic effects of other species. The Mayer method of the "infinite LC Zero" should be considered as an alternative to just using single acute data. Another important consideration is the effect that the characteristics of the water can have on toxicity. In the case of metals, softer water makes the chemical more toxic and turbidity mitigates the toxicity of lipophilic organic chemicals. These matters are not easy to include in a regulatory program. However, the Subcommittee challenges the Agency to make better use of current science. Defaulting to the statistical derived estimates with limited consideration for the complexity of water quality factors may not be serving the best interest of water quality.

3.2.2 Proposed Uses of the Tier 2 Values

The Subcommittee believes that Tier 2 values are compatible with the proposed uses, if used as a "value" and in a manner consistent with guidance very appropriately spelled out in the introduction of the EPA document (Host et al., 1990). The briefing of the Subcommittee by GLWQI personnel in February 1992, clearly indicated in the handouts the concept "Tier 1 numbers were to be adopted by Great Lakes states as numeric criteria" and Tier 2 "to be adopted as a narrative procedure". However, the Subcommittee is concerned about how Tier 2 will be implemented, particularly with respect to such issues as permits, permit limits, periods of time allowed for improving the data base, and anti-backsliding[1].

---

[1] Antibacksliding is a legal concept from the Clean Water Act which prohibits the relaxation of permit limits in some cases. EPA may allow such modifications under certain circumstances, which may include the existence of new scientific information which indicate that the underlying criteria are too stringent. Modification of permit limits may also be affected by anti-degradation provisions of state or federal law.

3.3   Major Issues Identified Related to GLWQI Tier 2 Approach

3.3.1 Tier 2 and the EPA Advisory Concept

The proposed GLWQI Tier 2 method is similar at least in its intent to the
Michigan Rule 57 and the U.S. EPA's "Guidelines for Preparing Water Quality
Advisories" developed several years ago. In a 1988 review of these guidelines for
advisories, the Environmental Effects, Transport and Fate Committee of the SAB
endorsed the aquatic life advisory concept while recognizing that advisories should
not be a substitute for continuing development of water quality criteria based on a
full set of data (Science Advisory Board, 1988). In addition, the Subcommittee
identified several issues that needed to be addressed in order to enhance the
potential utility of the approach. These included: the problem of implementation
where laws stipulate that state water quality standards cannot be made less
stringent, as would be the case as new data ultimately lead to a full water quality
criteria; a method to identify which chemicals deserved advisories; better
documentation of the uncertainty factors used; input data quality; inclusion of the
concept of exposure duration; and site specific modification possibilities.

In this review of the Tier 2 method, which was judged similar in intent and
use to the previous water quality advisory method, the SAB Subcommittee once
again endorsed the concept recognizing that many of the scientific deficiencies
identified in the previous method may have been addressed in the new statistical
procedure. However, most of the science-related policy and implementation issues
cited in the previous review are relevant.

3.3.2 Additional Testing

The Subcommittee noted that the second paragraph on page 1 of the
"Analysis of Acute and Chronic Data for Aquatic Life" that EPA realizes that
"although a criterion (full Tier 1)) might be desirable, it might not be necessary".
EPA further discussed how the Tier 2 data can be appropriately used to determine
whether a predicted or measured exposure concentration of a chemical in a body
of water is cause for concern because of toxicity to aquatic organisms. If the
margin of safety is sufficiently large Tier 2 data may be adequate without any
additional data. If the margin is small then there may be a justification to expect
additional toxicity testing in order to get better resolution on the safety issue. It
was recognized that this Tier 2 approach would help avoid the generation of
unneeded data. The Subcommittee feels that this is a good use of the Tier 2

15

approach and recommends its use in this manner. However, Tier 2 values should not replace the more scientifically defensible Tier 1 criteria.

3.3.3 Tier 2 Acute Factors and Acute/Chronic Ratios
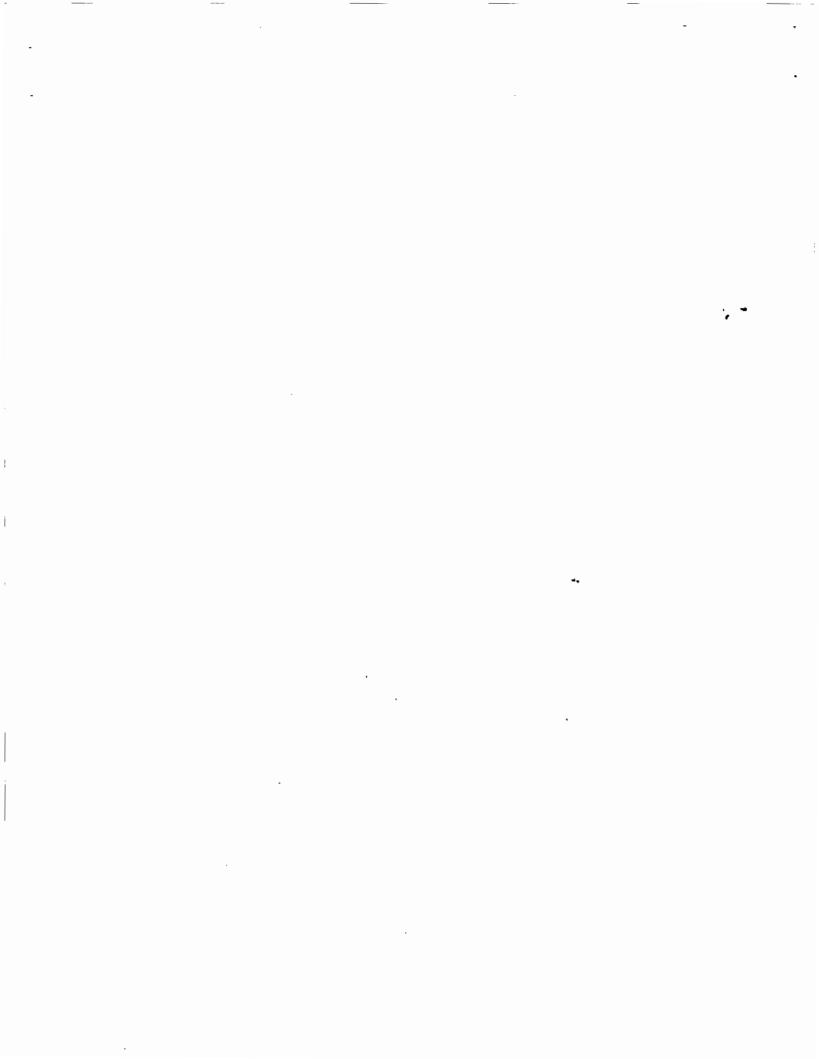
It was not possible for the members of this Subcommittee to judge the absolute validity of the statistical analysis of acute and chronic data. However, the method seemed conceptually correct but could benefit from a review by a separate group of experts in statistics. Many of the concerns expressed in the previous SAB review on the guidelines to derive water quality advisories appeared to be addressed with this new statistical method. However, the Subcommittee still cautions against misuse of the Tier 2 concept and specific values derived from the procedure.

3.3.4 Implementation of Tier 2 and Anti-Backsliding

The Subcommittee expressed concern on the issue of how the Tier 2 data might be used by some states and the implications of the current EPA policy of anti-backsliding. States implementing this Tier 2 method must realize that all Tier 2 estimates will, because of the statistical derivation method used, result in a value more stringent than a full criterion. As more data are obtained over time, the value will frequently become less stringent as it approaches the Tier 1 value. If these facts can not be dealt with in implementation then there can be no scientific defensibility in the Tier 2 concept.

3.3.5 Relationship of Tier 2 to Whole Effluent Toxicity

EPA recognized years ago the water quality criterion program (Tier 1) of the NPDES was not addressing all the needs in protecting the nation's water. Thus, the whole effluent toxicity testing program was developed and implemented to regulate the many unknown chemicals that were likely to cause adverse impacts in receiving waters. This whole effluent toxicity testing program is recognized by the Subcommittee as a valuable and scientifically justifiable program. It is somewhat redundant with the intentions of the GLI Tier 2 aquatic life criteria approach. The Subcommittee believes that if the Tier 2 program is implemented within the framework of the previous discussion on "guidance and anti-backsliding" then it can be a valuable additional tool in the hands of the water quality manager. It would be unfortunate if the Tier 2 method was used to generate

useless and unneeded information on long lists of chemicals; whole effluent toxicity testing should be included as an alternative.

The Subcommittee recommends that the GLWQI consider incorporating many of the practices and policies embodied in the "Whole Effluent Toxicity Tests" program. An "in-situ" fish or mussel bioaccumulation test or an HPLC test could be implemented, as part of a battery of tests, to alleviate the concerns of establishing limits based upon calculated bioaccumulation factors (BAF's) that may be either too stringent or inadequate to achieve the desired levels of protection.

The Subcommittee accepts the concept that a "field" measured residue of nonpolar contaminant relative to the mean "bioavailable" concentration in the water should be acceptable measure of BAF. Even this measure is subject to considerable error due to temporal changes in concentration of the contaminant, analytical errors associated with dissolved versus sorbed fractions, and uptake rates by individual fish. Also, the exposure time should be sufficient to allow for development of equilibrium conditions between the contaminant in the environment and the organism.

3.3.6 Site-Specific Variability

Criteria derived from the Tier 2 minimum species data set may not adequately consider site specific factors such as water hardness and warm versus cold water conditions. Application of such calculated values could result in unknown over- or under-estimation of the concentration needed to protect aquatic life. This can occur because fewer species tested under fewer water quality conditions can result in criteria with limited application over the range of water temperatures and hardness conditions of the Great Lakes Basin.

3.3.7 Laboratory to Field Validation

A need was recognized to review the protectiveness of these Tier 2 values in relation to real world impacts. The Subcommittee suggests that some of the existing field studies that have analyzed Tier 1 WQC for their application could be revisited with the intention of looking at the degree of conservatism that will result from Tier 2 values.

17

3.3.8 The WQC Data Base

The overall process of combining the acute and chronic toxicity data for the set of 29 chemical contaminants may have introduced artifacts when comparing toxic effects of metals, insecticides, solvents, petroleum hydrocarbons, etc. The mechanism of action of the different classes of contaminants could certainly influence the ratio of acute to chronic toxicity values. Thus, limits derived from the grouped contaminants could be either over- or under-protective of aquatic life exposed to a specific contaminant. All data were put into a single data set to increase the size of the database. Chemical effects data were not separated according to modes of action or obvious classes such as metal, pesticides, and others. It would have been more scientifically sound if this had been done, especially for the acute to chronic ratio. However, the reduced size of the data base would have reduced the robustness of the statistical parameters. As more data are accumulated over time, EPA should split the data as suggested to improve the quality of the estimates.

4. WILDLIFE CRITERIA

4.1 Introduction

The development of criteria for protection of wildlife is probably the most innovative aspect of the Great Lakes Initiative. The lack of such criteria has been a significant obstacle for the Agency with respect to its overall mission of protection of the environment. Although the Subcommittee has major reservations on the scientific defensibility of certain aspects of the present formulations of the wildlife criteria methodology and specific aspects of the proposed criteria, the Subcommittee wishes to encourage and support the Agency's efforts in the development of criteria for protection of wildlife. It should also be noted that habitat and disease may have major influences on the success of local populations of wildlife.

The development of methodologies for the criteria to protect human health and aquatic life has required considerable effort. The experience gained in developing these methodologies gives the developers of the methodologies for the derivation of criteria for the protection of wildlife a considerable head start. However, this does not imply that wildlife criteria can be generated through minor fine-tuning of existing criteria, or that the existing data base is adequate to generate criteria for all substances of concern without the need for further research.

4.2 Problem and Definition of Significant Terms

What is the definition of wildlife in the context of the criteria? Does the term refer to all animal species that are not domesticated, or does it refer only to air-breathing vertebrate species that are legally hunted, does it include invertebrates, or is it some intermediate definition? In the Great Lakes Initiative wildlife appears to have been defined in terms of a restricted number of piscivorous species: otters, mink, bald eagle, osprey, and kingfisher. These species occupy the apices of the water based food webs in the Great Lakes and would thus be expected to be highly exposed. The draft should explain that the representative species are not intended to be the most sensitive species exposed to the chemicals. Alternatively, the recent National Wildlife Criteria Methodologies meeting of the U.S. EPA in Charlottesville, Virginia (April 13 - 16, 1992) defined wildlife as mammals, birds, reptiles and amphibians. It is acknowledged that there is a body

19

of knowledge on the experimental biology of some amphibians which would permit the development of a toxicological data base in short order. However, in spite of this laboratory capability for studies in amphibians, such a toxicological data base has not been assembled. Beyond this, the basic knowledge on the natural history of reptiles is so fragmentary that it is not possible to maintain them routinely in the laboratory over a complete life cycle. Consequently it is not even possible to establish a toxicological data base covering full life cycles for representative reptiles at this time. Although the definition of wildlife given at the National Wildlife Criteria Methodologies meeting is broad, and even though the toxicological data base to support it is fragmentary, the broad definition of wildlife is more supportable than the limited list of species used by the Great Lakes Initiative.

Several important questions exist regarding the establishment of wildlife criteria. Should there be national criteria, regional criteria, aquatic wildlife criteria, or species specific criteria? Should the methods be developed in such a way that they are suitable for the development of criteria at either the national level, or at site specific levels, or at organism specific levels? If the methodology is properly constructed, it may be feasible to fulfill all of these roles. In light of these concerns the Subcommittee recommends that the Agency develop a general definition of "wildlife" and justify inclusions and exclusions of particular groups of animals.

4.3 Exposure Assessment

The Great Lakes are unique in their considerable geographical extent and the long residence times of water and persistent contaminants within the lakes. This aspect of the Great Lakes requires an understanding of the environmental transport and fate of the contaminants, and developing a basin-wide approach to their control. Fish and wildlife in the Great Lakes (and elsewhere) exhibit high body burdens of substances that tend to bioaccumulate. For some substances (DDT, dieldrin, PCBs) the body burdens in the recent past have exceeded those of today. High body burdens of some contaminants have been associated with adverse effects in field studies, although it remains controversial whether these associations are causally related. EPA has appropriately identified bioaccumulative chemicals and potential effects on wildlife as major issues of concern. However, little foundation was presented to indicate that the Great Lakes system is unique with respect to either how chemicals bioaccumulate or the inherent sensitivity of the species that reside in the Great Lakes basin and how their populations are exposed and at what level they will be protected.

20

4.4 Dietary and Drinking Water as Routes of Exposure

At present the exposure assessments for wildlife are based primarily upon bioconcentrated substances in foods plus direct uptakes of substances from drinking water. Given the emphasis on chemicals with high BAFs, the influence of the drinking water route of exposure is negligible for those specific chemicals. Furthermore, in the present Great Lakes Initiative, wildlife exposures via inhalation or dermal contact are not considered. These routes of exposure can become important for chemicals with significant vapor pressure and intermediate molecular weights.
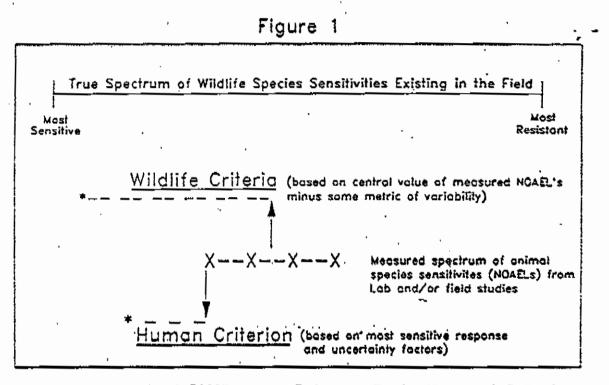
Overall, the proposed wildlife criteria methodology is confounded by combining exposure assessments (in terms of the BAF) with risk assessments (in terms of assessments of dose-response relationships extrapolated to daily intakes that do not produce adverse effects). If the Great Lakes Initiative's intent is to protect populations of wildlife, then it is important to control the daily absorbed dose of the chemical of concern to the members of the population of that species, regardless of the route of exposure. The BAF issue is significant, it needs to be explored on its own merits, but it should not confound the risk assessments in the development of criteria for the protection of wildlife.

4.5 Use of Representative Avian and Mammalian Species

Many species of mammals, birds, reptiles and amphibians are subject to substantial exposure to waterborne contaminants in the Great Lakes basin. An initial listing of all these species, with basic information on size, diet, and foraging zone, would show that the five piscivorous species considered in the document are not representative of the full range of any of these three characteristics. Additional species subject to substantial exposure include the raccoon, horned grebe, double-crested cormorant, green-backed heron, old squaw, black tern, common tern, Forester's tern, piping plover, tree swallow, snapping turtle and northern banded water snake. It is not self-evident that these species are "represented" by any of the five piscivorous species considered in the present document. In conclusion, the Subcommittee is concerned that the current approach does not adequately consider ecologically representative species in the selection of surrogate wildlife species. The Subcommittee recommends that if the agency chooses to address a specified list of species, then that list should be re-evaluated regularly and a rationale provided to add particular species.

21

Most of the proposed methodology for wildlife criteria is devoted to methods that estimate exposures. It is possible to estimate (with uncertainty) the total daily intake of a contaminant by an individual by estimating bioaccumulation in the food web, proportions of food arising from various trophic levels within that food web, daily food intake on the basis of allometric equations, and estimating the uptakes from other routes of exposure (inhalation and dermal). However, the methods do not consider the exposures from all environmental media. In addition, the estimates of daily intake are species and life-stage specific, they do not adequately address the question of the extent to which one species may be more sensitive than another, given the same amount of daily exposure. These differences in species sensitivities are currently addressed by the application of a Species Sensitivity Factor (SSF) which can vary from 0.1 to 1 (see later discussion on page 24).

The proposed GLWQI wildlife methodology is basically a modification of the methods to derive risk assessments on human health for substances that exhibit thresholds for non-cancer toxicological effects. Risk assessments for the protection of human health use information generated during the studies of several species to draw conclusions with respect to a single species - namely humans. Unless there is information to the contrary, it is assumed that humans are at least as sensitive as the most sensitive test in the most sensitive species tested. Furthermore, risk assessments for the protection of humans seek to protect the individual against effects which may be subtle, effects which relate to the quality of life rather than survival, and effects which may only occur in sensitive sub-groups and occur over long periods. Therefore, the tests that are incorporated into protocols applicable to developing criteria for the protection of human health can be very sensitive, and often go beyond the basic needs for protection of a species to assure the maintenance of its local population level. In contrast, criteria for the protection of wildlife must make allowances for differences in species sensitivities that are not incorporated into the methodology that has been developed for the protection of human health. When one seeks to protect a broad range of wildlife species, the challenge is to extrapolate from experimental data developed within a very limited group of laboratory species to the potential effects that may occur in the broad range of species whose populations need to be protected in the environment. This is complicated because the range of species sensitivities is not a constant, and the range of species sensitivities cannot be adequately captured in the allometric equations, which are largely related to species differences in dietary intake and body size. Independently of body size and dietary intake, the range of species sensitivities can be very narrow (e.g., with HCN or CO), or it can be very large

22

(e.g., with 2,3,7,8-TCDD). As illustrated below in Figure 1, there is a significant difference between the derivation of criteria for the protection of human health (lower arrow), and the relationship between laboratory studies and the needs to protect a broad range of wildlife species at the local population level (upper arrow).

Figure 1

```
┌─────────────────────────────────────────────────────────────────┐
│  ┌ True Spectrum of Wildlife Species Sensitivities Existing in the Field ┐  │
│                                                                   │
│  Most                                                       Most  │
│  Sensitive                                                  Resistant │
│                                                                   │
│         Wildlife Criteria (based on central value of measured NOAEL's │
│  *─ ── ── ── ── ──                  minus some metric of variability) │
│                        ↑                                          │
│                        │                                         │
│              X ── X ── X ── X .   Measured spectrum of animal     │
│                        │          species sensitivites (NOAELs) from │
│                        │          Lab and/or field studies        │
│                        ↓                                          │
│           * ─ ─ ─                                                 │
│           Human Criterion (based on most sensitive response       │
│                            and uncertainty factors)               │
└─────────────────────────────────────────────────────────────────┘
```

TITLE: Conceptual Differences Between Derivations of Criteria for Protection of Human Health and Wildlife

The question of how to account for the spectrum of differences in species sensitivities is far from simple. It is essentially impossible to identify the "most sensitive" species, because the most sensitive species status is likely to be chemical specific. Ideally one would have chronic toxicity data on a broad range of species, so that the range of species sensitivities could be determined directly.

Alternatively, one could determine the range of species sensitivities by some means of extrapolation based upon a statistical analysis of a sampling of data relevant to species sensitivities. Potentially useful examples are: the lower 95th confidence limit of the geometric mean of NOAELs from chronic toxicity studies in a spectrum of species; the 5th percentile of the percentile distribution of NOAELs from chronic toxicity studies; a chosen percentile of a Monte Carlo simulation from chronic NOAELs. If there are insufficient chronic toxicity studies, then one could resort to the range of species sensitivities in acute toxicity tests, coupled with the application of an acute toxicity to chronic toxicity ratio akin to the procedures employed for the derivation of ambient water quality criteria for the protection of aquatic life. However, the acute to chronic ratio is probably not a constant, as has been pointed out previously in SAB comments on the methodologies for the development of criteria for the protection of aquatic life. The Subcommittee recommends that the methodology for deriving wildlife criteria incorporate procedures that address a measure of the variability of species sensitivities observed in substance specific studies.

The proposed approach suggests that mink or kingfishers are the most exposed and/or sensitive species, and that a species sensitivity factor (SSF) ranging from 0.1 to 1 as a multiplier can account for any additional contingencies, considering that the exact value of the SSF needs to be based upon best professional judgment. As the procedures move from a direct assessment to indirect indicators, their reliability deteriorates.

Wildlife criteria need to take account of the principles and uncertainties of extrapolating information across evolutionary, spatial and temporal dimensions. Furthermore, the developers of the wildlife criteria methods also need to recognize that the uncertainties in this process accumulate in complex ways because the available data are based on conditions that may not occur in the field.

In special cases wildlife criteria need to be constructed so that they are able to protect the individual rather than the population. This can be an important consideration for endangered species, or for species covered by various treaty obligations that prohibit the "taking" of individuals.

4.6 Interpretation of Toxicity Data

The majority of the toxicological information for chemicals that are known to predominate in the Great Lakes system, has been derived either in direct

24

support of criteria for the protection of human health, or it has been generated as part of research into the basic toxicology of these chemicals. Consequently, much of the information that is cited in support of the wildlife criteria was generated in studies that were not designed to support the development of such criteria. Therefore, most of the extant information suffers from various deficiencies.

4.6.1 LOAEL vs. NOAEL

The discussions of the LOAEL v. NOAEL issues in the Technical Support Document are very superficial. The discussions cover the degree of adjustment required to estimate an NOAEL when the available information is based only upon an LOAEL. This issue is seriously confounded with the range of species sensitivities discussed above, and furthermore it is strongly influenced by the dosage spacing that was used in the chronic toxicity study. It is very important to remember that the entire evaluative structure involving the "NOEL - NOAEL - LOAEL - Severity of Effect" concepts, was formulated for a human health risk assessment paradigm. Although these concepts can serve as sources of inspiration, they cannot serve as foundations for the development of criteria methodologies for the protection of wildlife or anything other than humans. For example, biochemical changes and the induction of enzymes without concomitant histopathological changes, may be of significance for the development of criteria for the protection of the human individual. It is not at all apparent to what extent such changes might influence the long-term success of local or regional wildlife populations. The Subcommittee recommends that the Agency develop guidance for the selection of NOAELs appropriate for the protection of wildlife populations as distinct from the protection of individuals.

The principal message here is that the interpretations concerning specific effects extrapolated to the well-being of a human individual or those for the maintenance of sensitive wildlife populations, are fundamentally different. Consequently, it is unlikely that criteria for the protection of wildlife can be created by relatively minor adjustments to the methods that have been developed for the derivation of the criteria for the protection of human health. The uncertainty factor that seeks to relate the LOAEL to the NOAEL is related to the spacing of the dosing regime chosen by the investigator.

4.6.2 Subchronic to Chronic Extrapolation

The proposed methodology suggests that an up to 10-fold uncertainty factor be applied to subchronic studies. Reproductive and developmental toxicity studies are inherently short-term studies which do not merit the application of sub-chronic to chronic uncertainty factors. Even among target organ toxicity studies, there are many instances where sub-chronic studies are actually more sensitive than the chronic studies carried out on the same substance. As animals age, various organ systems deteriorate in function and histological structure (e.g., kidney studies by Coleman et al., 1977). This deterioration is common in most organ systems. As a result it becomes increasingly difficult to discern damage to these organ systems induced by chronic exposures to chemicals as animals age, because the deteriorating status of these organ systems in control animals obscures the effects produced by the exposure to the test substance. Further evidence for this phenomenon is provided by McNamara (1976) who found that 90 day studies were more sensitive than life-time studies in over half of all cases, so that the ratios between dose rates at which comparable effects were seen at 90 days relative to their dose rates at the end of the life span, ranged from 0.1 to 10. The methodology should discuss this problem and provide a rationale for when the 10-fold uncertainty factor is appropriate for subchronic results. Further, the methodology should note that a 10-fold uncertainty may also be appropriate for chronic studies due to the masking of effects caused by aging test animals.

4.6.3 Field and Laboratory Study Information

Experimental toxicology studies and field studies provide complementary information. Experimental studies provide precise dose-response information under simplified and controlled conditions. The causal relationship between dose and response can usually be clearly demonstrated, but the ability to interpret the applicability of the information is constrained by the artificiality of the test conditions.

Although field studies provide direct information on the response of species under real-world conditions, this information often exists only in terms of associations. Such associations offer important opportunities to identify effects of concern and target them for priority research. However, in field studies the question of causality can rarely be established without question. Hill (1965) listed a set of criteria for establishing causality when positive associations are found to exist (Table II).

Table II

Criteria for the Evaluation of Causal Associations

STRENGTH: a high magnitude of effect is associated with exposure to the stressor.

CONSISTENCY: the association is repeatedly observed under different circumstances.

SPECIFICITY: the effect is diagnostic of a stressor.

TEMPORALITY: the stressor precedes the effect in time.

PRESENCE OF A BIOLOGICAL GRADIENT: a positive correlation exists between the stressor and the response.

A PLAUSIBLE MECHANISM OF ACTION: some understanding of the functional relationship between stressor and effect.

COHERENCE: the hypothesis does not conflict with knowledge of natural history and biology.

EXPERIMENTAL EVIDENCE: laboratory results which support a hypothesis.

ANALOGY: similar stressors cause similar responses.

After Hill (1965).

The difference is closely analogous to that between experimental toxicology studies and epidemiological studies in human health risk assessment. Hill's criteria need to be applied with care. Absence of information on some criterion only implies that causality cannot be established, not that it cannot exist. Both laboratory and field studies provide important information, both contribute to the weight of evidence, and both should be used for complete assessment of qualitative and quantitative responses.

4.6.4 Tissue Residues

Tissue residues in the target species can be used as indicators of exposure, and if the organisms have achieved equilibrium concentrations, then they can also be used as measures of exposure. When the basic modelling information is available, then the tissue concentrations can be integrated with physiologically based pharmacokinetic models (PB/PK models) and these can in turn be integrated with laboratory toxicity information arranged in biologically based dose/response

27

models (BBDR models). These models are at the cutting edge of toxicological research, and they are being investigated vigorously, because they appear to offer the best way of bridging the differences in toxicological responses among different experimental systems. The field of wildlife toxicology needs to be a prominent part of this effort.

At present, tissue residue data have often been the first indication that the inputs of chemicals into a large system like the Great Lakes Basin have exceeded the rates at which the chemicals can become biologically unavailable through a variety of processes.

4.7 Tier 2

More information is needed before the Tier 2 wildlife proposal can be evaluated completely. In principle, it appears reasonable to develop interim "values" for chemicals for which some data are available, but are insufficient to establish a Tier 1 wildlife criterion. However, the major differences in the minimum data requirements specified for Tier 1 criterion development compared to Tier 2 revolves around the use of sub-chronic v. chronic studies. Compared to the problems inherent in the development of Tier 1 criteria, the further distinctions introduced in the Tier 2 methodology are trivial.

There are clearly advantages to having a form of Tier 2 criteria or "interim" criteria for wildlife. The present proposal represents only minor differences to Tier 1, it does not make scientific justifications for the magnitude of the uncertainty factor that Tier 2 requires, and it is does not provide justifications or scientific advice on implementation needed for risk management that is consistent with the concepts of "anti-backsliding, non-degradation, zero-discharge, and virtual elimination of toxics" programs.

4.8 Individual Criteria Documents

Previous reviews of these substances for other criteria or health assessment documents have required the full time efforts of special review panels specifically constituted for each substance. Typically each review has taken more than one day. This SAB Subcommittee was not constituted to conduct compound specific reviews.

4.9 Toxicity Equivalent Factors (TEFs)

The TEF values have been developed to estimate the relative toxicities of PCDDs and PCDFs, with a recent interest to include appropriate PCB congeners. The major impetus for this development has been the concern for carcinogenicity. Issues related to the fundamental assessment of the toxicity of dioxins and selected dioxin-like compounds were reviewed by the Dioxin Ecotox Subcommittee (Science Advisory Board, 1992). A concern expressed by this Subcommittee is whether TEFs developed to assess relative carcinogenic potency are also applicable to assess effects on reproductive and developmental toxicity. Furthermore, it is unclear to what extent TEFs developed largely in mammalian systems are applicable to avian or other wildlife species.

5. BIOACCUMULATION FACTORS

5.1 General Comments

The GLWQI documents present a good explanation of bioaccumulation and the need to consider it rather than only bioconcentration in the establishment of management scenarios for the Great Lakes environment. In the absence of field derived data, the Initiative attempts to generate criteria for human health and wildlife based on bioaccumulation factors derived from perceived trophic levels, organism lipid estimates, and octanol-water partition coefficient (K_{ow}). These endeavors are admirable and the Subcommittee encourages EPA and the Great Lakes states to continue to explore these approaches and address some problems associated with them.

The Subcommittee finds that the BAF procedure is more advanced and scientifically credible than existing simple BCF procedures. The use of the BCF, Food Chain Multiplier (FCM), and BAF approach appear to be fundamentally sound. However, a major inconsistency exists between field data for some chemicals (Reinert, 1970) and the conceptual model of Thomann (1989) for food chain derived residues. Efforts should be devoted to clarifying and improving the documentation and the issues discussed below with a view to presenting a straight-forward procedure with associated estimates of confidence levels. It is the Subcommittee's opinion that with some modification a credible BAF estimation method can be developed exploiting present knowledge.

5.2 Field Measured Bioaccumulation Factors

A "field" BAF is the ratio of the concentration of a chemical in feral fish to its concentration in water from the same locality. Generally the water concentration in question is "total" rather than truly dissolved or "available". Few such "field" data exist (see for example, Reinert (1970) and Reinert and Bergman (1974)), but they do demonstrate convincingly that field BAFs exceed laboratory bioconcentration factors (BCFs) by a substantial factor for many hydrophobic chemicals. While field measurements should be an acceptable measure of BAF, there can be considerable error due to factors such as temporal changes in concentration of the contaminant, analytical errors, whether dissolved or suspended concentrations were determined, variable uptake rates by individual fish, mortality of target species, and fish mobility.

Reinert (1970) and Reinert and Bergman (1974) found that the concentrations of DDT and Dieldrin in relation to fish lipids ("oils") were nearly constant across all aquatic trophic levels. Generally, the percent lipid increased at higher trophic levels and with the length of the fish. For these lipophilic pesticides, reporting residues based on lipid minimizes the effect of the food chain. EPA should update its model in relation to these data.

Field BAFs must be interpreted very carefully, and it should be recognized that they may contain substantial errors and variability due to the following reasons:

a)   Analytical methodologies generally determine total concentrations all of which may not be biologically available;

b)   There may be a loss of analyte by sorption or evaporation during sampling;

c)   Incomplete extractions may occur, especially if there is a high organic carbon content in the water;

d)   Temporal and spatial variability in water concentration may occur due to season, temperature, depth, hydrology, meteorology, and microbial and photolytic activity;

e)   There is likely to be variability in fish concentrations due to size, age, sex, season, pre- or post-spawning status, migration, the nature of and availability of food, the structure of the food chain, differences in lipid content, parasite infestation and general health of the organism.

Given these potentials for error, EPA should discuss and quantify the variance in field derived BAFs in its guidance, along with FCM estimates and attempt to identify the magnitudes of natural variability and analytical errors in each criterion data base, and estimate the impacts on the BCFs and FCMs.

In many cases, the laboratory generated BCF data are likely to be more analytically accurate, but they may be less representative than BAF, in that they do not reflect natural variabilities, especially on food uptake. Therefore, field measured BAFs are suitable for the calculation of criteria but with the qualifications that the data must be interpreted carefully and all information

31

should be exploited. Specific guidelines need to be developed for the acceptability of residue data in tissues and dissolved concentrations in water. This will likely require a research effort to determine the appropriate sampling procedures, such as the number of organisms per station, the sampling frequency, or filtered/ unfiltered water.

To help alleviate the problem, EPA needs to support a research program to develop more sensitive analytical methodologies for hydrophobic chemicals in tissues, sediments and water. Consideration should be given to the establishment of a formalized analytical chemistry program which utilizes the best scientists, the best instrumentation, adequate support, etc., to develop analytical methodologies' and perform analyses that are not readily achievable by "normal" laboratories. Support to universities and industrial support to develop analytical reference materials would help ensure the success of the program.

At present, water concentrations in field derived bioaccumulation calculations are assumed to be totals, i.e., $C_T=C_W+C_P$. There are abundant arguments in literature that show that dissolved (C_W) and particle bound (C_P) contaminants have different availabilities over time. This concept is certainly recognized in the Agency's effort designed to develop sediment quality criteria. One can thus ask the question: if the science underlying the development of sediment quality criteria recognizes partitioning because dissolved and particulate associated contaminants present different bioavailabilities, why do water quality criteria not incorporate this state-of-the-art understanding of speciation and phase partitioning? The Subcommittee recommends that these factors be presented as part of the criterion methodology with a clear and defensible explanation as to why GLI ignores these factors.

5.3 Adjusting BCFs to BAFs

Theoretically derived bioaccumulation factors appear to be based upon accepted concepts of how chemical exchange between water, food, and fish; but they have not been applied to enough field conditions to judge if the predictions are realistic. Thomann's (1989) model for bioaccumulation incorporates the appropriate transfer coefficients for uptake via food intake and allows for rates of excretion. Biotransformation can be included, however rates of biotransformation cannot be estimated adequately from physical/chemical properties such as K_{ow} and therefore must be determined experimentally for each compound, or at least each functionally related group of compounds. There is also considerable uncertainty

32

about the factors controlling food uptake efficiency. The Thomann (1989) model assumes that the lipid-normalized BCF is equal to K_{ow} at zero growth and at "equilibrium". This basic assumption does not allow for oxidative metabolism and biosynthetic conjugation with hydrophilic ligands such as glucuronic acid, sulfates, and acetates. The model has not been adequately tested to use for the establishment of regional water quality criteria at this time. The potential exists for errors on both over-protection and under-protection of aquatic organisms, wildlife and humans. It is noteworthy that almost all bioaccumulation work has focussed on non-metabolizing, non-polar, chlorinated hydrocarbons. Relatively little has been done on metabolizable chemicals such as PAHs or phenols.

The Subcommittee is particularly concerned that consideration of metabolism is not included. Admittedly it is difficult to find rate constant data but for certain chemicals such as the PAHs, however, metabolism is an important determinant of BAF. Metabolism may become more significant when lipid stores are reduced at times of stress and lipophilic chemicals become mobilized. Metabolism is also an important detoxification mechanism. In principle, metabolism can invalidate the use of the simple FCM approach but the Subcommittee is unable to suggest an alternative other than the use of reliable field BAF's.

It should be noted that Thomann's (1989) model gives only very general expressions for respiration rate, feeding rate and growth rate as a function of organism mass. More accurate species-specific data exist for these rates which could be used instead, presumably giving greater accuracy. The option to use such data should be included.

At present the GLI procedures use an equation for BCF developed by the Duluth Environmental Research Laboratory plus the Thomann (1989) equation for FCM. The Subcommittee recommends that the GLWQI use either the entire Thomann (1989) approach, which has been tested or test the validity of the GLWQI combination of approaches. The significant difference is that the Veith and Kosian approach does not view the bioconcentration as simple lipid partitioning.

Laboratory generated BCF values can be measured in a number of ways. Systems prescribed by EPA and OECD include: static, sequential static, semi-static, and flow-through systems. In addition, conditions such as times of exposure and kinetic frameworks may be specified. It is now becoming evident that Log

K_{ow}-BCF relationships work well for chlorinated organic chemicals with low to medium molecular weights (<500 - 600) and for which no biotransformation occurs. Carefully specified procedures for measuring and estimating BCFs and Log K_{ow} for other classes of compounds must be developed and evaluated. A proper testing protocol should be able to accommodate questions as to such effects as: influence of pH, especially for those compounds that dissociate; the influence of mixtures on bioavailability, solubility, and general partitioning; the influence of age on different fishes and their capacity to bioaccumulate; and the influence of a third phase (i.e. suspended or bottom sediments) on BCFs.

BCF relationships for metals present a special problem, as recognized by the authors of the GLI document. The Subcommittee strongly urges the authors to pay particular attention to the fact that total analytical concentrations of metals (and organically complexed metals) may not represent the "activity" of that metal. Enough is known now about aqueous metal speciation, precipitation behavior, and solids partitioning to incorporate this body of knowledge into scientifically rigorous criteria protocols. The Subcommittee recommends that the GLWQI collaborate with modeling specialists from the EPA Athens laboratory.

5.4 BCF for log K_{ow} above 5.0

At present the BAF confidence intervals for chemicals with log K_{ow} < 5 appear to be quite tight while those in the range of 5 to 6.5 have confidence intervals which may be more than an order of magnitude wide. In the range beyond 6.5, the confidence is not known within reasonable limits. This situation is less than satisfactory for a regulatory program.

The treatment of super hydrophobic chemicals, e.g., those with log K_{ow} > 6.5, by assigning them an arbitrary FCM of 1.0 is viewed as merely an admission of ignorance. This presents a problem in that most of the chemicals in this range have high molecular weights and volumes and they may be subject to slow absorption and clearance as a result of retarded diffusion through absorbing tissues. EPA should consider other approaches to handling these substances such as:

a) Using only field BAF's in such cases;
b) Conducting chemical specific assessments;
c) Assuming all chemicals with log K_{ow} > 6.5 behave similarly to one with a log K_{ow} = 6.5 for which the BCF is accurately known.

34

This is clearly an area in which more research is needed. The field of bioaccumulation is clearly of importance in the GLI process. Considerable progress has been made in recent years towards understanding the factors influencing BAFs. Notable are the experimental studies at the EPA laboratory in Duluth and the modeling work at the Athens laboratory and at Manhattan College. There is a need to bring together the available scientific expertise in water chemistry analysis, fish physiology and pharmacokinetics, biochemistry, and food chain structure and fish ecology with mathematical modeling to derive credible, validated BAF models. A combination of thoughtfully designed laboratory studies and field investigations is needed. To date, the work being exploited is fragmented. The modeling approach of the Athens group is promising, but there is a need for more active cooperation between modelers and biologists, the latter being in the best position to understand the nature of the complex series of events which comprise bioaccumulation. In short, EPA should mount a specific research program in this area to satisfy the needs of programs such as the GLI.

5.5 Analytical Methodology for Compliance

The present GLI document presents numerical criteria values for four chemicals. The criteria for several of these chemicals, and presumably for a host of others, will be less than the analytical detection capabilities of many laboratories. Additionally, because aquatic systems receive contaminants from other point and non-point sources, the ambient concentration of a chemical in question may already exceed proposed criteria. The Subcommittee recognizes that it is entirely plausible to arrive at a scientifically defensible concentration value which is orders of magnitude below present analytical detection levels. However, EPA should provide implementation guidelines for the discharger to deal with monitoring criterion values that are below present limits of detection and situations where present ambient concentrations, derived from possibly a multitude of sources (including atmospheric sources), far exceed proposed criteria.

The Subcommittee suggests several approaches. The dischargers could estimate the mass of discharged pollutant and, together with known volumes of water flow, calculate theoretical concentrations; if the substance is hydrophobic, samples of the suspended solids should be analyzed. It is possible that analysis of surficial sediments maybe more reliable than analysis of water. The dissolved and total concentrations could be estimated using partitioning theories. Caged fish could be used as "accumulators" for specified time periods to estimate BCF values

for field exposures. Surrogate "accumulators", such as dialysis bags filled with appropriate solvents could also be used.

6. HUMAN HEALTH CRITERIA

6.1 General Comments

Conceptually, the Great Lakes Initiative has significant implications for improving the ability to assess health hazards associated with water contaminants in the Great Lakes Basin. The National Program has never before specified a minimum data set for estimating water quality criteria related to human health. The tiered approach suggested offers a mechanism for improving the data base necessary to reduce the uncertainties in risk assessment in the national criteria' and to develop appropriate data for compounds released to the Great Lakes Basin for which national criteria do not exist. On the other hand, the process has the potential for being somewhat frivolously applied to chemicals which should be regarded as safe without specific testing. Within this group would fall natural substrates consumed in food as a source of calories (e.g., fatty acids, sugars, amino acids).

There are also some serious difficulties with the way the tiers are constructed. It is not possible to argue that Tier 1 chemicals protect against reproductive developmental/teratogenic or carcinogenic endpoints because the minimum data base does not require data that are appropriate for estimating hazards of these types. In other words, chemicals can be classified as non-carcinogens or without reproductive/developmental effects by not being tested for these endpoints. The lack of such data is adjusted for by additional uncertainty factors in the case of reproductive/developmental effects, but no adjustment was made for lack of data on carcinogenesis.

The Subcommittee suggests that compounds that lack data on carcinogenesis, reproductive and developmental/teratogenic effects be relegated to Tier 2. Therefore, Tier 2 would serve to identify those chemicals for which an inadequate data base exists. The generation of the appropriate data could be rewarded by smaller uncertainty factors and movement into a Tier 1 criteria.

Certain of the criteria developed in the Great Lakes Water Quality Initiative will inevitably bring up conflicts in risks to the ecosystem or specific wildlife relative to risks to human health. A specific example of such a circumstance would be the disinfection of wastewater effluents to prevent infectious disease transmission. Such treatments inevitably lead to the formation of a variety of by-

37

products which may represent some finite ecological risk. To reduce or eliminate disinfection of such effluents might result in a substantial impact on human health. The potential magnitude of such risks relative to risks associated with some potential bioaccumulative compounds whose estimated risks frequently depend upon multiplicative factors that may or may not be realized in the real world might present a very distorted view of the benefits of regulation. The documentation provided does not provide any perspective to how such issues might be resolved.

The Subcommittee suggests that the EPA should be much more explicit in balancing human health risks with ecological risks. The EPA should either exempt chemicals added to water for public health purposes or modify ambient water quality criteria to allow for prudent use of these chemicals. This needs to be addressed at the national level, not only in relationship with the GLWQI.

Another major concern is what the impact this local or regional activity will have on the development of national standards. While many of the aims of this program are laudable, differences in the criteria or the types of compounds that are limited between the national and local level are likely to lead to confusion and distrust on the part of the public served.

6.2 Thresholds for Carcinogens

The approach taken in carcinogenesis risk assessment in the Initiative is over simplistic and does not break out the questions in such a way as to encourage concise development of the rationale for the risk assessment. The current technical guidance should be revised to reflect the following discussion.

Weight of evidence issues simply address the question of whether available data indicates that humans would be sensitive to the carcinogenic effects of the chemical. This decision would be based on the consistency of the carcinogenic response across species using criteria articulated by both EPA and IARC. A more recent refinement of these criteria includes specific questions of whether the mechanism for producing cancer in experimental animals exists in humans. This latter question has been explicitly considered in recent deliberations on chemicals that induce accumulation of alpha-2U-globulin in male rats, compounds which are peroxisome proliferators in rodents and substances which induce thyroid tumors.

38

The method by which low dose extrapolation is conducted should not be simply viewed as a question of threshold or non-threshold carcinogens. To the extent possible the carcinogenic response should be modeled in the context of the mechanisms by which the chemical induces the cancer. These considerations need to be included independently of whether the chemical has been shown carcinogenic in humans or only in experimental animals. Several research groups are developing data that will allow the independent contribution of mutagenic and cell proliferative effects of a chemical to the carcinogenic response to be modeled. Additionally, data are being developed which allow the effective dose of the responsible metabolites to be estimated across species with much more confidence utilizing physiologically-based pharmacokinetic (PBPK) models. When such information is available it should be utilized in estimating carcinogenic risks at low doses. As it happens both mechanistic and pharmacokinetic data are being developed for a number of compounds that are being considered for developing water quality criteria. It is important to note that treatment of individual chemicals in this process may be complicated and perhaps controversial. The Subcommittee recommends that the GLWQI coordinate with other programs in EPA that are currently addressing these issues. A partial list of chemicals that will receive much attention in the near future are dioxin, chloroform, PCBs, trichloroethylene, tetrachloroethylene, and the phthalates.

In the interim, the linearized multistage model is a reasonable default methodology for the many chemicals for which these more detailed data are not available. However, the Agency must utilize available information on the mechanisms carcinogens to modify this assumption whenever data are considered to be reliable.

6.3 Additive Risks for Carcinogens

The assumption of additivity for chemical carcinogens is difficult to accept as a default at low doses. Additivity assumes a common mechanism of action. This is probably an infrequent condition since compounds classified as carcinogens are known to act by a wide variety of mechanisms and to target different organs. Moreover, compounds which operate by different mechanisms, mutagenic versus non-mutagenic, are likely to be synergistic at effective doses but less than additive at low doses.

Within the confines of compounds that act at the same receptor (e.g., dioxins, furans and PCBs) an assumption of additivity might well be defensible.

However, it must be recognized that such interactions can also be antagonistic, a weaker activator of the receptor may actually inhibit the effects of a more effective activator (i.e. acting as a partial antagonist). On the other hand, synergistic activity can be expected to occur between compounds acting by different mechanisms within the same target organ. The classic example are compounds that act by mutation vs. induction of cell proliferation. It is clear, however, that the latter chemicals must be presented in doses that either produce sustained levels of increased cell division or prevent cell death in the target organ. Thus, such interactions are unlikely to occur at low doses.

The SAB recommends that the GLI consider the probability of interaction between carcinogens on a case by case basis. These interactions must also be taken into account within the context of their co-occurrence in fish tissue rather than from simple projections of their concentrations based on occurrence in effluents. The compounds might well take entirely separate environmental pathways. It would be unwise to project potential errors of an interaction on top of errors in risk assessment and projections of bioaccumulation.

6.4 Tier 1 Minimum Data Base

Tier 1 should be reserved for those compounds that have been adequately tested. To include chemicals in Tier 1 which have not been adequately tested for carcinogenic, reproductive or developmental/teratogenic effects is inconsistent with the stated goals of the initiative. A corollary to this is that a Tier 1 compound should not carry an extra uncertainty factor for lack of appropriate data.

With this suggested reconstitution of Tier 1, the C carcinogen classification would include chemicals that had been adequately tested and the weight of evidence does not support the notion that they are probable human carcinogens. Therefore, they would not be treated as carcinogens in developing the criteria. These chemicals should be clearly distinguished from those which have received their C classification because they have only been tested in a single species. The absence of data in one or more other species essentially frustrates the development of a weight of evidence argument. Therefore, such chemicals should be relegated to Tier 2 and treated as suspect carcinogens. A more conservative assessment could be justified to force completion of a data base on the chemical that will allow a proper judgement to be made.

6.5 Tier 2 Concept

Under the proposed Tier 1 and Tier 2, carcinogens classified as C would be treated more conservatively if they fell into Tier 2. This seems more consistent with the overall aims of Tier 2, the development of a guidance that encourages the development of more definitive data.

The use of a 28-day study (the term subacute should be abandoned) to produce a NOAEL as the minimum data set for Tier 2 may be marginal for detecting some chronic human health effects. The latent period for some well defined chronic effects approach this limit (e.g., peripheral neuropathies produced by acrylamide, n-hexane or methylbutylketone) and some exceed it (e.g., peripheral neuropathy produced by dichloroacetate). More speculative neurotoxicities may have even longer latencies (e.g., aluminum induced neurofibrillary tangles).

On the other hand, data from experiments of 28 days duration are better than no data at all. It is important that these data be developed in an accepted mammalian species and that additional uncertainty factors compensate for the significantly less sensitive detection limits that will result from the shorter duration of the experiments.

A more broadly cast function for the Tier 2 concept should make it more defensible. As indicated above one may reluctantly start with a minimum data base of a 28-day study, but other critical data deficiencies should also place compounds into this class

6.6 Chemical Mutagens

There are substantive difficulties in determining whether chemicals induce these effects by "genotoxic" mechanisms. Consequently, it is impossible to consider this question without knowing what the minimum data base for determining whether a genotoxic mechanism is involved. Clearly one cannot use bacterial or *in vitro* methods for determining whether a chemical will produce a heritable mutation in humans. Damage to the germ cell line can only be gaged by *in vivo* heritable mutation assays. On the other hand, the whole animal tests available in this area are inordinately insensitive or very expensive to conduct.

In addition to demonstrating that the effect is the result of a genotoxic mechanism, some effort must be expended to develop a methodology for

41

extrapolating these data to low doses. Risk assessment models to predict the induction of heritable mutations and/or genotoxic teratogenic effects must consider other factors (e.g., spontaneous abortion) that influence the outcome. Thus, simple application of the same risk assessment models used for cancer may not be appropriate. Consequently, a case must be developed for any extrapolation model that would be used before the Subcommittee can usefully comment on this proposal. There have been some recent efforts in this area that have been published in the literature, but consensus on this issue has not yet been reached.

6.7 Relative Source Contribution

Most persistent and bioaccumulative environmental contaminants offer a substantial exposure potential through the food chain. In the case of water contaminants, fish flesh represents a predominant food source of exposure. The procedure proposes a default 80% relative source contribution (RSC) for water contaminants to account for fish flesh exposure as a predominant human exposure. There might be individuals that have very high exposures in contaminated soil, but this is such an irregular source that development of an RSC is unlikely to protect such individuals anyway. Other sources (i.e., fish taken from other sites) are really already compensated for in the calculation of fish consumption. Consequently, there seems little justification for any RSC for these chemicals. The SAB believes that a factor of 80% is not supportable because it is within the rounding error on the calculations of the overall exposure.

6.8 Additional Concerns

a) The criteria documents must explicitly consider issues that are critical to the development of the either Tier 1 or Tier 2 criteria. Examples are: 1) It is important to consider both positive and negative carcinogenesis data to develop a weight of evidence argument that a chemical is or is not to be considered a carcinogen. 2) The selection of a NOAEL is strengthened by finding similar values for similar endpoints in several independent experiments. 3) The methods of quantitative extrapolation of risks need to be supported by data relevant to the chemical's mechanism(s) of action when such data is available.

b)   Risk Assessments do not consider multiple routes of exposure posed
     by volatile compounds.

          The basis of the criterion is an assessment of human exposure
     by drinking water and fish consumption only. Exposure by air
     inhalation of chemical derived from water is not considered and
     should be. For example, the toluene drinking water criteria of 6000
     μg/L or 6 mg/L corresponds to an equilibrium concentration in air of
     about 2 mg/L or 2 g/m^3. This greatly exceeds the occupational health
     TLV (188 mg/m^3). Water containing these concentrations of toluene
     will taste and smell offensive. Additionally, there is the fact that
     levels of 22,000 μg/L or 22 mg/L could represent a fire hazard or an
     explosion hazard in sewer systems. The point is that water and fish
     consumption are not the only exposure routes or hazards to be
     considered, especially for volatile chemicals.

          EPA should consider adding an air inhalation term to the
     exposure denominator in the form of some volume of air V_a inhaled
     which achieves equilibrium with the water. It would become W_c +
     (FC x BAF) + V_a x K_{aw}, where K_{aw} is the air-water partition
     coefficient.

c)   The data utilized and assumptions made to arrive at the 15 g/d for
     consumption need to be more explicitly discussed.

d)   Data and assumptions used to arrive at the lipid content need to be
     made explicitly in the document.

Literature Cited

Coleman, G.L., Barthold, S.W., Osbaldiston, G.W., Foster, S.J. and A.M. Jonas. 1977. Pathological changes during aging in barrier-reared Fischer 344 male rates. J. Gerontology 32(3):258-278.

Hill, A.B. 1965. The environment and disease: association or causation? Proceedings of the Royal Society of Medicine 58:295-300.

Host, G., Regal, R. and C.L. Stephan. 1990. Analyses of Acute and Chronic Data for Aquatic Life. EPA Environmental Research Laboratory-Duluth.

Lebo, J.A., Zajicek, J.L., Huckins, J.N., Petty, J.D. and P.H. Peterman. 1992. Use of semipermeable membrane devices for insula monitoring of polycyclic aromatic hydrocarbons in aquatic environments. Chemosphere 25:697-718.

Johnson, G.D. 1991. Hexane filled dialysis-bags for monitoring organic contaminants in water. J. Environ. Sci. Technol. 25:1897-1903.

McNamara, B.P. 1976. Concepts in health evaluation of commercial and industrial chemicals. Chapter 4 in: New Concepts in Safety Evaluation. M.A. Mehlman, R.E. Shapiro, H. Blumenthal, eds. Advances in Modern Toxicology 1(1):61-140. Hemisphere Publishers, Washington, D.C.

Reinert, R. 1970. Pesticide concentrations in Great Lakes fish. Pesticide Monitoring Journal 3(4):97-111.

Reinert, R.E. and H.L. Bergman. 1974. Residues of DDT in Lake Trout (Salvelinus namaycush) and Coho Salmon (Oncorhynchus kisutch) from the Great Lakes. J. Fish. Res. Board Can. 31:191-199.

Science Advisory Board. 1988. Review of the Guidelines for Preparing Water Quality Advisories. EPA-SAB-EETFC-88-032, July 1988, Washington, DC.

Science Advisory Board. 1992. Review of EPA's Research to Support Water Quality Criteria for TCDD (Dioxin). EPA-SAB-EPEC-92-024, August 1992. Washington, DC.

Thomann, R.V. 1989. Bioaccumulation model of organic chemical distribution in aquatic food chains. J. Environ. Sci. Technol. 23:699-707.

ABBREVIATIONS AND DEFINITIONS

BAF   Bioaccumulation Factor, a ratio of the concentration of a chemical in fish or tissue to concentration of the chemical in water.

BCF   Bioconcentration Factor, chemical-specific values used to predict tissue residues derived through direct uptake of the chemical from water.

Board   See SAB.

C_{18}   Radioactively labelled carbon.

DWC   Drinking Water Committee, a standing committee of the Science Advisory Board.

EHC   Environmental Health Committee, a standing committee of the Science Advisory Board.

EPEC   Ecological Processes and Effects Committee, a standing committee of the Science Advisory Board.

FAV   Final Acute Value, a numerical estimate of the concentration of a chemical in water that will protect aquatic life from acute toxicity. This is usually a Tier 1 value based on a minimum data set of toxicity test on a variety of aquatic species.

FCV   Food chain multiplier, a value used to account for the concentration of residues predators obtain from prey.

FWS   U.S. Fish and Wildlife Service.

GLI   Great Lakes Initiative.

GLWQI   Great Lakes Water Quality Initiative, a coordinate EPA and State effort to protect aquatic organisms and humans from the adverse effects of persistent toxic pollutants in the Great Lakes.

GMAV   Genus Mean Acute Value, another value used in the calculation of aquatic life water quality criteria.

46

HPLC	High Pressure Liquid Chromatography, a tool in analytical chemistry used to measure the quantity of organic compounds.
LC	Lethal Concentration, the concentration of a chemical which kills a proportion of the test organisms (e.g., an LC50 is the concentration that kills 50% of the test organisms).
LOAEL	Lowest Observed Adverse Effect Level, the lowest test concentration of a chemical at which a deleterious effect was measured.
NOAEL	No Observed Adverse Effect Level, the highest test concentration of a chemical at which no deleterious effect was measured.
NPDES	National Pollution Discharge Elimination System.
PAH	Polyaromatic Hydrocarbon compounds (e.g., styrene).
SAB	Science Advisory Board, a public advisory group operated by staff of the EPA Administrator.
SAV	Secondary Acute Value, a Tier 2 value based on a limited data set, designed to protect aquatic life from acute toxicity.
TEF	Toxicity Equivalent Factor, a system for comparing the potency or toxicity of mixtures of congeners of a chemical.
Tier 1	Criteria values which are based on a minimum data set of toxicity testing and bioaccumulation data for a chemical. The nature of the minimum data set varies among aquatic life, wildlife, and human health criteria.
Tier 2	Criteria values based on a limited data set of toxicity testing for a chemical.
WQC	Water Quality Criteria, numerical estimates of the levels of a chemical in water that will protect aquatic organisms and humans. These values are used by states as the foundation for state water quality standard which are used to set limits in discharge permits.