UNITED STATES ENVIRONMENTAL PROTECTION AGENCY
WASHINGTON, D.C. 20460

January 9, 1992

OFFICE OF
THE ADMINISTRATOR

EPA-SAB-EPEC-92-014

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

RE: SAB Report "Review of a Testing Manual for Evaluation of Dredged Material Proposed for Ocean Disposal"

Dear Mr. Reilly:

The Sediment Criteria Subcommittee of the Ecological Processes and Effects Committee of the Science Advisory Board has completed its review of the draft manual "Evaluation of Dredged Materials Proposed for Ocean Disposal - Testing Manual". This manual, commonly referred to as the "Green Book", is an update of the 1977 version. It provides National technical guidance for use in establishing the Limiting Permissible Concentration (LPC) for ocean disposal of dredged material. The manual uses a tiered approach, which the Agency's Contaminated Sediment Strategy Task Force is also considering for use in other programs.

The Subcommittee was asked by the Office of Water to review a portion of the 1991 Green Book to assess the adequacy of the bioaccumulation and toxicity testing procedures in the manual and to identify other methods which could address chronic toxicity and bioaccumulation. The Subcommittee met on April 15-16, 1991 and again on September 24, 1991 to complete the review.

The Subcommittee found that the revised manual is a significant improvement over the preexisting 1977 "Green Book", with respect to the use of the reference sediments, the selection of test organisms, and the option of using a 28-day bioaccumulation study. There are, however, several changes which should be made now, including those which: identify which tests and test organisms are mandatory; standardize definitions of endpoints and stressors; and provide an

interpretative framework for decisions at the local Corps of Engineers and EPA offices. On a somewhat longer time scale, the Subcommittee recommends that EPA make additional changes to: focus the tiered evaluation process in the manual toward reducing uncertainty, develop guidelines to interpret bioaccumulation tests, conduct studies of the relationship of body burdens to "no observed effect levels", update the concept of an application factor, and define how sediment quality criteria will apply to the "Green Book".

The Sediment Criteria Subcommittee has been asked by the Office of Water to conduct another review of the technical guidance needs for disposal of dredged materials at inland sites that are covered under the Clean Water Act. Several of the recommendations from this review should also be considered by EPA in the development of this second guidance.

We look forward to receiving your response to our recommendations and we are particularly interested in any plans that the Agency may have to further its understanding of the evaluation and interpretation of bioaccumulation of substances from sediments and the use of application factors. We look forward to advising you in the future on these and other issues related to contaminated sediments.

Sincerely yours,

Raymond Loehr, PhD.
Chair
Science Advisory Board

Kenneth L. Dickson, PhD.
Chair
Ecological Processes and
Effects Committee

Robert Huggett, PhD.
Chair
Sediment Criteria Subcommittee

EPA

United States
Environmental
Protection Agency

Science Advisory
Board (A-101)

EPA-SAB-EPEC-92-014
January 1992

AN SAB REPORT: REVIEW OF A TESTING MANUAL FOR EVALUATION OF DREDGED MATERIAL PROPOSED FOR OCEAN DISPOSAL

PREPARED BY THE SEDIMENT CRITERIA SUBCOMMITTEE OF THE ECOLOGICAL PROCESSES AND EFFECTS COMMITTEE

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 does 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 do not constitute a recommendation for use.

U.S. ENVIRONMENTAL PROTECTION AGENCY
SCIENCE ADVISORY BOARD
ECOLOGICAL PROCESSES AND EFFECTS COMMITTEE
SEDIMENT CRITERIA SUBCOMMITTEE

ROSTER

CHAIRMAN

Dr. Robert Huggett, Virginia Institute of Marine Science, School of Marine Sciences, College of William and Mary, Gloucester, Virginia

MEMBERS

Dr. William Adams, ABC Laboratories, Columbia, Missouri

Dr. Donald F. Boesch, Director, Center for Environmental and Estuarine Studies System, University of Maryland, Cambridge, Maryland

Dr. Peter Chapman, E.V.S. Consultants, Vancouver, British Columbia, Canada

Dr. Rolf Hartung, School of Public Health, University of Michigan, Ann Arbor, Michigan

Dr. Frederick K. Pfaender, Department of Environmental Sciences and Engineering, University of North Carolina, Chapel Hill, North Carolina

FEDERAL AGENCY LIAISONS

Dr. Robert Scott Carr, NFCR Field Research Station, US FWS, Corpus Christi, Texas

Dr. Suzanne Bricker, NOAA, Rockville, Maryland

SCIENCE ADVISORY STAFF

Dr. Edward S. Bender, Biologist and Designated Federal Official

Ms. Marcia K. Jolly, Staff Secretary

Mr. Robert Flaak, Asistant Staff Director

Dr. Donald G. Barnes, Staff Director

ABSTRACT

This report represents the findings and recommendations of the U.S. Environmental Protection Agency's Science Advisory Board summarizing a review of the "Evaluation of Dredged Materials for Ocean Disposal-Testing Manual". The review was conducted by the Sediment Criteria Subcommittee at two meetings on April 16-17 and September 24, 1991. The manual outlines a tiered testing approach for evaluating dredged materials for compliance with the limiting permissible concentration (LPC) as defined by the Ocean Dumping Regulations. The Subcommittee reviewed the adequacy of the bioaccumulation and toxicity testing procedures in the manual and provided recommendations for mandatory tests, selection of test organisms, and the development of a regulatory framework for interpreting the results. The Subcommittee also recommended that EPA revise the guidance to clarify the use of the tiered approach, to elaborate the requirements for evaluation under tier IV of the scheme, define the relationship of Sediment Quality Criteria to the "Green Book", expand guidance on the selection of reference sites, and to address several scientific questions about bioaccumulation and its effects on the organism. The Subcommittee further recommended several editorial changes for clarity and consistency of the definitions, eliminating redundancies, and consolidating the quality assurance requirements for the test results.

Key Words: Dredged Materials; Bioaccumulation Potential, Toxicity; Tier testing.

TABLE OF CONTENTS

1.0 EXECUTIVE SUMMARY 1

2.0 INTRODUCTION 2
2.1 Statement of the Charge 2
2.2 Subcommittee Review Procedures 3

3.0 EVALUATION OF THE BIOASSAY AND BIOACCUMULATION METHODS FOR DREDGED MATERIALS 4
3.1 General Comments 4
3.1.1 Tiered Testing 4
3.1.2 Microbiological Concerns 5
3.1.3 Quality Assurance and Quality Control 5
3.1.4 Definitions 6
3.1.5 Synergistic Effects 6
3.1.6 Use of Sediment Quality Criteria 7
3.1.7 Redundancy 7
3.2 Selection of Organisms used in Bioassays 7
3.2.1 Recommended Species List 7
3.2.2 Criteria for species selection 8
3.2.3 Bioaccumulation Tests 8
3.2.4 Development of Tests 8
3.2.5 Mandatory Tests 8
3.3 Bioassays and Endpoints 9
3.3.1 Water Column Concentrations 9
3.3.2 Application Factor 10
3.3.2 Theoretical Bioaccumulation Potential 12
3.3.3 Rationale for Bioaccumulation Tests 13
3.3.5 Toxicity Test Endpoints 14
3.4 Reference Area Concept 15
3.5 Interpretative Guidance 15
3.6 Tier IV Requirements 16

4.0 SUMMARY OF RECOMMENDATIONS 17

5.0 LITERATURE CITED 19

1.0 EXECUTIVE SUMMARY

The Sediment Criteria Subcommittee met twice April 16-17 and September 24, 1991, to review the bioassays and bioaccumulation test procedures that were used in the draft manual entitled, "Evaluation of Dredged Materials for Proposed Ocean Disposal-Testing Manual (February, 1991)". This manual is referred to as the "Green Book" and represents an update of a 1977 version. Both manuals were developed jointly by the Corps of Engineers and EPA. The Subcommittee was asked to evaluate the adequacy of the bioaccumulation and toxicity test procedures in the revised manual for assessing contamination of sediments in the National Ocean Disposal Program.

The Subcommittee found that the revised manual represented a significant improvement over the 1977 version with regard to the use of the reference sediments, the selection of test organisms, and the option of using a 28-day bioaccumulation study. In its deliberations, the Subcommittee found that this manual addressed several principles that had significant implications for the evaluation of dredged materials for disposal in freshwater as well as for managing contaminated sediments in other situations. However, the Subcommittee believed that the manual could benefit from careful editing and further refinement of the tiered approach.

There were several changes, which the Subcommittee has recommended, that should be made now, including those which: specify which tests and test organisms are mandatory; standardize definitions of endpoints and stressors; and provide an interpretative framework for decisions at the local Corps of Engineers and EPA offices.

On a somewhat longer time scale, the Subcommittee recommends that EPA focus the tiered evaluation process toward reducing uncertainty, develop guidelines to interpret the results of bioaccumulation tests, conduct studies of the relationship of body burdens to "no observed effect levels", and re-evaluate and modify, if necessary the concept of a constant application factor. The tiered approach should also be revised to allow the reevaluation at higher tiers for sediments that failed at a lower tier. The manual should be revised to address microbiological issues or renamed and provide the reader with other sources or references for test procedures to address these issues. Guidance on quality assurance should be consolidated and a glossary of terms should be added and used in a consistent fashion throughout the document. The Subcommittee recommended that EPA define the relationship between the "Green Book" and Sediment Quality Criteria.

2.0 INTRODUCTION

The maintenance of navigation channels requires dredging. The disposal of that dredged material is a concern to government, industry, and public sectors. For ocean disposal, the dredged material must be evaluated to determine its impact adjacent to the disposal site. In 1977, EPA and the Army Corps of Engineers (COE) developed a manual "Ecological Evaluation of Proposed Discharge of Dredged Material into Ocean Waters" which contains technical guidance on chemical, physical, and biological procedures to determine the acceptability of dredged material for ocean disposal. In February 1991, the EPA and COE published a revised draft of this manual, commonly called "the Green Book" that included more testing procedures. This manual will be updated in the future in accordance with revised rules for Ocean Dumping.

The manual outlines what is called a tiered testing procedure for evaluating compliance with the limiting permissible concentration (LPC) for long term exposures that are defined by the ocean dumping regulations. The liquid phase or water column LPC must not exceed applicable marine water quality criteria or a toxicity threshold (0.01, an application factor, times the acutely toxic concentration). The suspended and solid phase LPCs must not cause unreasonable toxicity or bioaccumulation. The manual describes four levels (tiers) of evaluations. Tiers I and II utilize existing information, which is often available for recurring disposals of dredged materials from channel maintenance to determine the appropriateness of ocean disposal. Tier III contains most of the laboratory bioassays, and Tier IV includes some tests of bioaccumulation and a range of possible field investigations. The manual also recommends use of a reference site, which is intended to be free from contamination, as source of sediments for comparison testing with the dredged materials.

2.1 Statement of the Charge

On March 6, 1991, the Office of Marine and Estuarine Protection requested that the SAB review the toxicity testing and bioaccumulation methods of the "Green Book". As part of that review the Subcommittee was asked to address the following question:

> Are the bioassays, (bioaccumulation, chronic/sublethal, and acute toxicity tests) currently in use or under development by EPA and the Corps of Engineers, appropriate for assessing contamination of sediments for use in the National Ocean Disposal Program? How can they be improved?

The Subcommittee accepted the charge but noted that it was also necessary to review the bioassays in the context in which they would be applied. In particular, the Subcommittee felt the review must include some aspects of the tiered approach and the use of a reference site to evaluate some of the bioassay results.

2.2 Subcommittee Review Procedures

The Sediment Criteria Subcommittee, a standing group of the Ecological Processes and Effects Committee of the SAB, was established to review and evaluate methods and procedures for assessing the quality of sediments. The membership of the Subcommittee has changed to provide expertise that is needed for specific reviews, however, the Chairman and a small nucleus of experienced members have been retained to provide consistency. For this review, the Subcommittee relied on background materials and briefings on the manual and the development of methods of testing acute and chronic toxicity.

The Subcommittee met twice to review the Dredged Materials guidance, first on April 16-17, 1991 and again on September 24, 1991 in public meetings. Between the two meetings, draft comments were prepared and EPA reconsidered the charge for the Subcommittee as discussed in section 2.1. The Subcommittee expects to address broader aspects of sediment control strategies in another review in 1992. During the September meeting, the Subcommittee was also advised of additional guidance that EPA and the Corps of Engineers are developing for disposal of dredged materials in freshwater and estuarine sites under the jurisdiction of the Clean Water Act. Following the public meeting, the Subcommittee held a writing session prepare the revise the draft report.

3.0 EVALUATION OF THE BIOASSAY AND BIOACCUMULATION METHODS FOR DREDGED MATERIALS

3.1 General Comments

The document represents an improvement over the preexisting 1977 "Green Book". It includes a hierarchial testing approach with emphasis on the use of aquatic toxicity tests and it provides better guidance on species selection and use of reference material. It also provides the option of using a 28-day bioaccumulation study. The approach, although not incorporating a true tiered approach, should help eliminate unnecessary testing as well as identify those sediments that should be tested further.

The emphasis of the manual is clearly on toxicity and bioaccumulation test methods, with brief coverage of physical and chemical characterization of sediments and dredged materials. Microbiological issues, introduction of new species, genotoxicity testing, etc., are not covered. If EPA wishes to maintain the manual's focus on toxicity and bioaccumulation testing, it should consider renaming the document to better describe its contents or include a section in future drafts identifying all the testing required by the legislation and directing the user to other documents that contain testing procedures for parameters not covered by this document.

3.1.1 Tiered Testing

The Subcommittee supports the concept of tiers, best professional judgement, site-specific information, expertise and judgement calls. However, the framework for their use and the limitations to such use must be clearly enunciated. Conceptually, increasing tiers should correspond to decreasing uncertainty. This is not the case in the present document. Tier III appears always to be required. Additive effects are always expected to occur and are the basis of EPA's own risk assessment guidelines for mixtures (U.S. EPA, 1982). Synergism, if defined as a greater than additive response, cannot be predicted for complex mixtures. It is to be expected in an unknown proportion of cases and can only be established by determining the toxicities of all of the ingredients of the mixture compared to the toxicity of the mixture itself (Hartung, 1988; Stara & Erdreich, 1985; N.R.C., 1988). Consequently Tier III, as presently defined, will always be required. Also, there is no connection between Tiers III and IV. If applicants fail Tier III, there is no alternative to prove that the sediment is acceptable and, similarly, regulators also have no mechanism to require further testing if sediments pass Tier III.

Most tiered assessment systems are designed so that the lower and more simplistic tiers embody more conservative assumptions than do the higher tiers. This is the correct strategy, because the minimal information requirements in the lower tiers yield a broader range of uncertainties than the more extensive information requirements in higher tiers, However, this strategy produces the

ramification that some dredged materials that would fail a lower tier, could pass when tested in a higher tier. Unless the criteria for passing the higher tier requirements are fatally flawed, such a dredged material should be perfectly suitable for disposal. However, in the present strategy a material that fails at any tier is not eligible for more detailed evaluation at a higher tier. Further, the document clearly states that applicants can skip tiers. True tiered testing does not have superfluous tiers and does not allow entry at any point. True tiered testing allows for flexibility to challenge the assumptions and preliminary results at lower tiers; the present "tiers" minimize flexibility. **The Subcommittee recommends that EPA revise its tiered approach to allow reevaluation at higher tiers of failures at lower tiers, to clarify the linkages between tiers, and to emphasize reducing uncertainty as the level of tier testing increases.**

3.1.2 Microbiological Concerns

Sections 227.7(c)1, 228.10.6 and perhaps other portions of the revised rules for Ocean Dumping deal with the relocation of materials containing pathogenic microorganism, including viruses. This may be a very real issue, both in instances where sewage- contaminated materials are being dredged, especially if they are deposited on beaches, and for those human pathogens indigenous to the marine environment, i.e. *Vibrio parahaemolyticus* and *Vibrio vulnificans*. The testing manual contains no guidance on how dredged materials, control and reference sites should be evaluated for this class of biological contaminants. EPA's response to this issue seems to be that the manual under review deals with chemical toxicity and bioaccumulation testing and therefore microorganisms were not included. Since pathogenic microorganisms of fecal origin are usually present in environmental samples at concentrations too low to be directly detected by routinely available methods, we use bioassays of indicator organisms to provide evidence of their presence and the potential for human disease. Many different indicator organisms and testing approaches are available, ranging from the traditional coliform bacteria and fecal streptococci to more recent suggestions of bacteriophage and enteric viruses. The strengths and weaknesses of each have been extensively reviewed and discussed. The fact that there are significant problems in the application of some of these indicators and detection approaches in marine waters suggests that guidance is needed in a document that emphasizes bioassay testing methods. At a minimum the topic should be mentioned and the user directed to pertinent documents, many generated by EPA such as the National Estuary Program Monitoring Guidance Document, APHA Standard Methods for the Examination of Water and Wastewater, and EPA's methods published under Section 304.602 of the Clean Water Act that contain appropriate guidance and methods.

3.1.3 Quality Assurance and Quality Control

All quality assurance/quality control discussion should be consolidated in one section. Throughout the manual there is a great deal of redundancy regarding QA/QC and there is some incorrect information. For example, 24-h reference toxicant tests are recommended for amphipods, whereas 96-h testing is the correct

requirement. **The Subcommittee recommends that QA/QC be a separate, detailed Section, and that testing protocols be appropriate and clearly specified.**

Method Detection Limits should be provided for all chemicals in all media; in certain cases, there could be ranges rather than a single value. In either case, guidance is required and could be provided with a table showing how much tissue is required to achieve what detection limits by what analysis. This information would provide guidance for *a priori* selection of species and methods.

3.1.4 Definitions

Definitions are not always provided for critical terms used in the document, including: chronic, sublethal, sublethal chronic. The definition of bioaccumulation needs to be consistently applied throughout. The definition of LPC uses the wording "unreasonable toxicity", which should be defined. Three separate terms are defined under "may": may, can and might. These terms should be defined separately so the reader can find them. The terms "as similar as practicable" and "to the maximum extent possible" under Reference Sediment should be clarified. Preparation of Whole Sediment allows press-sieving and the "additions of small amounts of seawater"; the latter is of concern regarding dilution of the chemicals and the modification of their toxicity and no documentation is provided to support the statement "These procedures are unlikely to substantially alter chemical or toxicological properties of the respective whole sediments." "Pollution" as used on page 2-7, line 3 of paragraph 1 should also be defined. The document contains many abbreviations (e.g., LPC, LBP, TBP, MDL, etc.), but it does not contain a list of abbreviations used for reference purposes.

The Subcommittee recommends that a Glossary of Terms be generated and that the terms, as defined, be used in a consistent manner throughout the document.

3.1.5 Synergistic Effects

On pages 2-5, 2-6 and 3-11 the use of the words "synergistic effects" does not accurately reflect the state-of-the-art for aquatic toxicological tests. The draft document requires that the LPC be 0.01 of the acute toxicity when there are no water quality criteria or when the chemicals present are suspected of producing "synergistic effects". Years of acute toxicity testing and to a lesser extent chronic toxicity tests with aquatic organisms for a wide variety of test chemicals have indicated that synergistic effects are very rare. Most chemicals when tested in combination show additive effects or in some cases antagonistic effects. The manner in which the phrase "synergistic effects" is used in the document is either inappropriate or the authors meant "additivity".

The treatment of the toxicity of mixtures in the document does not take full advantage of the existing scientific literature. The NAS/NRC, U.S. EPA, and others have published reviews on the toxicity of mixtures which lend support to the

hypothesis that the acute toxicity of a mixture is, on the average, predicted by the sum of the acute toxicities of the components in the mixture. Thus, additive toxicity is the expected outcome, although greater than expected toxicities (variously defined as supra-additive, synergism, potentiation), as well as less-than expected toxicities (antagonism or infra-additive), have also been observed experimentally.

3.1.6 Use of Sediment Quality Criteria

EPA is in the process of developing guidance for the development and application of sediment quality criteria (SQC) based on equilibrium partitioning. The approach being used for non-ionic chemicals is to compare the interstitial water concentration (based on the organic carbon normalized whole sediment chemical concentrations and partitioning between the organic carbon and interstitial water) with existing water quality criteria. This methodology has been reviewed by the SAB ("Evaluation of the Equilibrium Partitioning (EqP) Approach for Assessing Sediment Quality" EPA-SAB-EPEC-90-006, February, 1990) and the reader is referred to this document for an evaluation of its strengths and weaknesses. Since it is not clear how SQC will be utilized in the disposal of dredged materials, **the Subcommittee recommends that the applicability of sediment quality criteria to dredged materials be delineated in the new "Green Book" Guidance.**

3.1.7 Redundancy

The Manual is repetitive in places, especially in the first seven Sections. In addition, there is a marked contrast in the level of guidance provided by equally important sections. **The Subcommittee recommends elimination of redundancies and repetition such that important information is clear and not obscured.**

3.2 Selection of Organisms used in Bioassays

3.2.1 Recommended Species List

Consistent with the objectives of the regulations, it is important that "appropriate sensitive" organisms be selected both for water-column and sediment bioassays. It should be understood that not all species deemed appropriate based on the criteria listed in Section 11.1.1 are necessarily sensitive and that the range of sensitivity of within a listed species is highly variable. Thus, it would be entirely possible to select test species of very low sensitivity and maximize the potential for passing the tests. However, it is also recognized that there are valid reasons to select species which are economically or ecologically important in the receiving waters, either to allay public concerns or to assess potential ecological effects in the field. It is important to include bioassays of species known to be highly sensitive to toxic effects of dredged material elutriates, including those designated by an asterisk in Table 11-1, sea urchin larvae to the veliger stage and bivalve larvae to the prossidonch stage. Additional bioassays of economically or ecologically important species which may not be highly sensitive may also be included, but should, to the extent practical, be based on the most sensitive life history stages of

these species. The use of sensitive species mentioned above will lessen the need for an application factor of 0.01 of the LC50 (Figure 3-2, page 3-9).

The Subcommittee recommends that Table 11-1 be amended such so that at least one species with an asterisk must be tested and adult bivalves should be removed from this list. The option of testing with sand dollar larvae should also be included. Further, larval tests should be clearly defined as involving development of the fertilized egg to the prossidonch I (bivalve) or veliger (echinoderm) stage.

3.2.2 Criteria for species selection

There must be a clear and consistent rationale and explanation for the species selection decisions. For instance, the argument for using zooplankton rather than phytoplankton (Page 2-8, second line of first paragraph) is that the latter are extremely variable. Zooplankton are also naturally very variable and in many cases populations increase and collapse in direct response to changes in phytoplankton populations. There are other cogent reasons for using zooplankton, including techniques available, sensitivity, use as indicators, etc., and these should be clearly stated.

3.2.3 Bioaccumulation Tests

The list of appropriate test species for bioaccumulation testing is provided in Table 12-1 and includes both organisms in intimate contact with the sediment and some that are not. **The Subcommittee recommends that Table 12-1 be amended such that recommended species be those in intimate contact with the sediments or those that ingest sediments.**

3.2.4 Development of Tests

The Army Corps of Engineers has devoted a considerable research effort to developing a "chronic" toxicity test with the polychaete *Neanthes arenaceodentata* for testing in Tier IV. However, evidence to date indicates that the *Neanthes arenaceodentata* 20-day growth and survival test is no more sensitive than acute lethality tests with infaunal amphipods. Since the purpose in recommending chronic tests in Tier IV is apparently to achieve increased sensitivity, it would appear that other test methods and test species other than *Neanthes arenaceodentata* may be better suited for this task. **Accordingly, the Subcommittee questions whether the effort and emphasis on the *Neanthes arenaceodentata* "chronic" test is warranted. We do not question a general emphasis on polychaetes and amphipods for test development, as long as the tests are appropriately sensitive and useful.**

3.2.5 Mandatory Tests

The Subcommittee agrees with the current emphasis on acute toxicity, specifically the limitation to mortality, and with the addition of the amphipod

toxicity test. However, at present, the amphipod test is a recommendation, not a requirement.

There are several specific requirements in the Ocean Dumping Regulations (Title 40, Code of Federal Regulations, Part 220-228) which have not been adequately addressed in the green book testing guidance. In Subpart G 227.27 (b), the regulations state "The limiting permissible concentration of the suspended particulate and solid phases of a material means that concentration which will not cause unreasonable acute or chronic toxicity or other sublethal adverse effects based on bioassay results using appropriate sensitive marine organisms." The guidance in the new green book does not provide recommendations for determining either chronic or sublethal effects in solid-phase toxicity tests.

Some species of infaunal amphipods are considerably more sensitive than the other "appropriate test species" endorsed for use in solid-phase tests in the new green book. Although we recognize that there is no single bioassay organism which is universally applicable for testing aquatic sediments, the guidance in the new manual could permit an applicant to essentially insure that no significant toxicity would occur by selecting the more insensitive "appropriate test species". **The Subcommittee recommends that the inclusion of a sensitive infaunal amphipod or preferable an appropriately sensitive test incorporating measures of chronic sublethal effects, as stated in the Ocean Dumping Regulations, be required in Tier III.**

The tests and test species recommended for solid-phase testing are not nearly as sensitive as those recommended for water-column testing and there is no "safety" factor or other means of providing a margin-of-error in the interpretation of the toxicity data from solid-phase tests. In addition, the reference sediment, which is collected from the vicinity of the disposal site, may be contaminated from previous disposal operations or by its proximity to other anthropogenic inputs and therefore serves to increase the probability of contaminated material being judged suitable for ocean disposal.

3.3 Bioassays and Endpoints

3.3.1 Water Column Concentrations

Page 5-2 of the document indicates that "step 1 of the tier II water column evaluation comprises a screen that assumes that all of the contaminants in the dredged material are released into the water column during the disposal operation." If this results in a concentration in the water column that exceeds a water quality criterion, then a leachate test is performed to better estimate the actual water column concentration. Scientific knowledge regarding the bioavailability of chemicals sorbed to sediments indicates that the presence of sediment reduces the bioavailability of chemicals as compared to aqueous exposure without the presence of sediment. Therefore, the above assumption is a very worst case and most often unrealistic. The results of the elutriate tests should provide a large data base to

draw from which would indicate the extent to which the bioavailability is reduced. These data could be used to provide a more accurate assumption for step 1 of the tier II water column evaluation.

3.3.2 Application Factor

The application factor is defined as:

AF = ["Safe" concentration]/[96hr. LC50 concentration]

Application factors were proposed as early as 1957 by Henderson to estimate "safe" concentrations for compounds for which no chronic toxicity data were available (NAS, 1972). It was recognized from the beginning that the AF was not a constant, but that it varied for different substances and species. The use of generic AF values for estimating "safe" concentrations, when only acute toxicity data were known, was proposed by the NAS/NAE Committee on Water Quality Criteria (NAS, 1972). This publication used "...present knowledge, experience and judgement..." to determine two classes of toxic compounds, for which different application factors were applied (page 121 of the document):

1. non-persistent (half-life <4d) or non-cumulative chemicals were considered "safe" at the following concentrations:

   0.1 X 96-h LC50 at any time or place
   0.05 X 96-h LC50 for a 24-h average

2. persistent, accumulative chemicals were considered "safe" at the following concentrations:

   0.05 X 96-h LC50 at any time or place
   0.01 X 96-h LC50 for a 24-h average.

Specific wording of the recommendation is provided on page 269 of the document:

> *In general, marine life with the exception of fish-eating birds and mammals should be protected where the maximum concentrations of the chemical in the water does not exceed one-hundredth (0.01) of the LC50 values listed in Column 7, table 10-7, pp. 265-268. This factor is arbitrary and was derived from data available on marine life and freshwater organisms (See Section III, p. 121). It assumes that a concentration of one-hundredth (0.01) of that causing harm to the most sensitive species to be protected will not damage this species or the ecosystem. Future studies may show that the application factor must be decreased or increased in magnitude.*

Where data were available, this document recommended specific application factors for the following chemicals, derived by dividing the "safe" concentration by the 96-h LC50:

Freshwater Application Factors

PAEs, Cu - 0.1
Cyanides, detergents, phenolics - 0.05

Marine Water Application Factors

Fluorides, cyanides, ammonia, chlorine, boron, sulfide - 0.1
Ag, Ba, Mo, Tl, Va - 0.05
Ni, Mn, Pb - 0.02
Al, Cu, As, Cr, phosphorous, Se, Cd, Be, Zn, uranium - 0.01

Clearly, the 0.01 application factor was a "worst case" estimate made almost two decades ago in the expectation that future research would result in refinement of values for specific chemicals. Some refinement was even provided in the document. Instead, the 0.01 value has become dogma. For instance, page 10 of the 1985 Technical Support Document for Water Quality-Based Toxics Control (September 1985, 74pp + appendices) notes:

> *State standards often include the application factors (to derive a "safe" toxicity value from an acute toxicity value) that were recommended in the 1972 "Blue Book" guidance (i.e., 0.05 for nonpersistent wastes and 0.01 for persistent wastes).*

A variant of this approach is used in the development of present water quality criteria, which employ the acute to chronic ratio (ACR) to estimate chronic concentrations from acute toxicity data. The ACR is not a generic value, but is calculated for each chemical on the basis of bioassay data.

The new "Green Book" (page 11-9, Section 11.1.7) refers to the LPC as being "...0.01 of the 48- or 96-h LC50, depending on the test duration." This is a very conservative estimate that assumes persistent chemicals are present; if nonpersistent contaminants are present, this application factor may be too stringent, particularly if truly sensitive test species are used, as we recommend. Further, this application factor was intended to apply to acute toxicity data (e.g., it does not apply to bivalve EC50 data).

Clearly, specific rather than generic application factors are justified by the original [19 year old] source document. **The Subcommittee recommends that EPA produce a peer-reviewed report that provides direction on the application factor issue appropriate for the 1990s and that the manual clearly indicate that the 0.01 application factor is only potentially appropriate for LC50 data with persistent chemicals.**

3.3.2 Theoretical Bioaccumulation Potential

The degree of bioaccumulation is considered to be unacceptable if animals exposed to dredged material bioaccumulate statistically greater amounts of contaminants than do animals exposed to reference sediments (page 2-3). This goal statement is reiterated in different words in section 5.2 (pages 5-5 to 6). The sections of the act are referenced, but the pertinent points of the act are not highlighted. The acceptability of the bioaccumulation is to be judged by the Theoretical Bioaccumulation Potential (TBP).

The basic formula that is proposed for calculating the Theoretical Bioaccumulation Potential (TBP) is:

$$TBP = 4\ (C_s/\%TOC)\%L$$

This simplistic relationship seeks to summarize the differences in the partitioning characteristics of sedimentary total organic carbon (TOC) and the total lipids present in the biota compared with the concentration in the reference sediment (Cs). The derivation of the proportionality factor of "4" encompasses many kinds of variability which become important in the assessment. The Committee recommends that the derivation of the factor "4" be documented to include the variance associated with this point estimate.

The assessment of the Theoretical Bioaccumulation Potential is largely dependent upon the comparative statistical distributions of the non-polar contaminants in the dredged material and at the reference site. Both the concentrations in reference sediments and dredged materials are expected to vary, and of course the concentrations of contaminants and total organic carbon concentrations are expected to co-vary. The overall null-hypothesis is that reference sediments are not significantly different from dredged materials. The rigor of this assessment is dependent upon the choice of the reference site and the adequacy of the sampling of the reference site and the dredge materials in relation to the occurrence of α and β errors in hypothesis testing. These complex relationships are not clearly stated in the manual. It is possible to discern some of these problems by careful analysis of topics that are widely dispersed in the manual.

At present, the extent of bioaccumulation is judged with respect to compliance with FDA guidelines and with respect to significant differences from reference sites. In the absence of FDA guidelines the judgment criteria can generate ambivalent results if the reference site is also contaminated. When the concentration of contaminants in the dredged materials exceeds the concentration found in the reference sediments, then the present methodology forces a decision that there is a problem, regardless of whether there is an effect that is discernible in biological studies. The manual indicates (pages 2-4 and 6-4) that a statistically significant increase in chemical concentrations in the tissue of organisms exposed to dredged materials as compared to organisms exposed to reference sediments is interpreted as a specific biological effect and has the potential to cause ecologically

unacceptable impact. The Subcommittee points out that statistically significant increases in tissue concentrations do not constitute biologic effects per se. The current guidance on bioaccumulation is driven by existing statutes. **It is recommended that appropriate consensus guidelines be drafted for interpretation of bioaccumulation tests and that these guidelines be incorporated into the regulatory framework. It is further recommended that the Agency conduct studies that relate body burdens to "no observed effect levels" (NOEC) as part of chronic toxicity protocols.**

As an adjunct to using the FDA guidelines, the Agency should routinely measure the concentrations of test materials in the test organisms at the termination of selected long-term toxicity studies. The important endpoint is the body burden of the chemical found at the "No Observed Effect Concentration" (NOEC). In long-term toxicity studies (on the order of months) the test organism should have achieved a practical bioconcentration equilibrium with the test substance, so that the concentration of the substance in the organism at the NOEC could be used just as an FDA guideline level. In the case of organic substances it may be useful to normalize the concentrations to the percent lipids in the organism. There is some evidence that the effective concentrations of specific chemicals at target sites are comparable over a wide range of species.

3.3.3 Rationale for Bioaccumulation Tests

The major concern over bioaccumulation is the build-up of toxicants in the food chain and, ultimately, human health risk. Thus, the major reason for this relatively expensive testing is to protect against biomagnification. But very few chemicals have been shown to biomagnify. Moreover, biomagnification is usually restricted to air breathing vertebrates (e.g., mink, heron, gulls, raptors, and seals). Finally, there are no good endpoints for bioaccumulation testing at this time.

The blanket recommendation on the duration of 10 or 28 days for the uptake studies is not appropriate. The basic problem with a 10-28 day bioconcentration test is that the bioconcentration coefficient is defined as the ratio of the concentration found in the exposed organism as related to the concentration of the contaminant in water, determined at equilibrium. There is no assurance that the exposed organism has achieved equilibrium at the end of 10 of 28 days. This problem cannot be solved by inspection of the data, but can only be solved by applying pharmacokinetic modeling to the test data (Spacie & Hamelink, 1985). However, if bioaccumulation testing is necessary at lower tiers due to regulations (it is not clear that this is, in fact, the case), then the 28-d test is a reasonable compromise at present between over stressing organisms and getting a clear signal. Still, it is important to note that steady state is not reached in all cases and the accumulation factor of 4, based on PCB's, does not provide any additional information. The range of mean application factor can be up to 20X different as was apparent from presentations to the Committee from the COE. Further, the statement that mortality of up to 25% in 28-d tests is acceptable (page 13-33,

narrative line 14) is confusing given previous, apparently conflicting guidance on acceptable mortality in these tests.

The suggested duration of exposure is clearly judgmental and arbitrary. However, all of the evidence that has been uncovered to date supports the fact that bioaccumulation in aquatic organisms follows a subset of the same rules that govern the discipline of pharmacokinetics. Additional confounding factors for poikilotherms and long durations of exposure are 1.- the influence of temperature on uptake and clearance rate constants, and 2.- the influence of growth during prolonged exposures. Regardless of these factors, the time to approximate equilibrium concentration is governed by an inverse exponential function of the clearance excretion rate constant K in the form of $F(1-e^{-Kt})$. Consequently the time to approach equilibrium is some multiple of the clearance (or depuration) half-life. In practice it can be very difficult to determine, on the basis of inspection alone, when an equilibrium concentration has been reached. Furthermore, clearance half-lives for compounds with high K_{ow} values (octanol to water partition coefficients) are often longer than 30 days and can approach a year. **The Subcommittee recommends that the estimation of the equilibrium concentrations be based upon pharmacokinetic modeling.**

3.3.5 Toxicity Test Endpoints

In certain situations where additional information is necessary for a thorough assessment of the biological effects component of the testing (i.e., Tier IV), it may be appropriate to test for sublethal effects using sensitive species. Observations of survival, growth and/or reproduction should not be abandoned. However, the Agency and the Corps of Engineers should address the potential for genotoxic, immunological, histopathological and other biochemical and physiological effects which are not presently addressed by the testing guidance document. There are a number of tests with marine organisms currently available to determine the potential for these types of effects. These non-conventional end points should be included as part of tier 4 in the "other sublethal effects" that should be considered when determining the limiting permissible concentration (LPC), as stated in the Ocean Dumping Regulations.

The responses of an organism to exposures to a given chemical are influenced by a wide range of extrinsic and intrinsic factors. The intrinsic factors include, for example, species, sex, life-stage, age, individual variation, genetic differences in enzyme levels, etc. The extrinsic factors include disease history, nutritional status, ambient physical and chemical environment, etc. The combination of these intrinsic and extrinsic factors exert strong influences on the variability of responses to toxicological insults. The impacts of specific factors are likely to differ when one compares environmental exposures with laboratory tests, and exact equivalence among responses under these two conditions is unlikely. The certain existence of these complications demands great care of the developer of testing protocols.

3.4 Reference Area Concept

While we agree with the reference area concept, reference areas must be better defined and quantified (as must control areas). The collection of reference sediment is not mentioned. Terms such as "as similar in grain size to dredged material as practical" must be defined. Reference areas must represent test sites as closely as practical.

The usage of the reference area concept is not clear or consistent throughout the new "Green Book". For instance (page 1-6, Reference Sediment), it is stated that if adequate reference sediments cannot be obtained, organisms not sensitive to grain size should be used in testing. This statement is repeated in the last paragraph on page 3-1 and top of page 11-10. However, later (page 11-1, 3rd paragraph, lines 11-12) an alternative is stated: "...or the effects of grain size have to be determined and considered when designing benthic bioassays and evaluating the test results." The text must clarify what is and what is not an alternative, and must be consistent.

The decision to permit disposal of dredged materials is to a large extent determined by the results of either toxicity or bioaccumulation tests performed with the dredged materials and reference site sediment samples. The selection of appropriate reference sites and the number of samples taken are critical factors in the final decision relative to the acceptability for ocean disposal of dredged materials. It is recommended that the reference site and sample selection guidance provided in pages 1-6 and 3-2 be expanded. Additional guidance is needed for determining when the "reference-point approach" should be used as compared to the "reference area approach". This guidance should discuss the needs for an appropriate statistical design that accounts for the variance that exists in the reference samples and the bioassay comparisons with the dredged materials. The statistical design should provide guidance for setting both the alpha and beta Type I and II error levels, the number of samples needed, and the conditions under which samples may or may not be combined for testing or analysis. Additionally, the Subcommittee recommends that the Ocean Disposal Manual prohibit the use of reference sediments from the proposed disposal area except where this area has not previously been used for disposal of contaminated material.

3.5 Interpretative Guidance

Interpretative guidance is lacking in the document; each region is allowed to proceed independently. This lack of guidance is not satisfactory as it is likely that different regions will make different decisions and there will be no consistency. Present provision of guidance (page 1-2, paragraph 6, lines 3-5) does not include comprehensive information on uses, limitations, and interpretation. In some cases (e.g., Tier IV), guidance on "...sound interpretation of the results..." is deferred to local agencies. The Subcommittee recommends that interpretative guidance, in the form of a framework for action, be provided now. This guidance should also provide the user with a list of critical changes that may be appropriate relative to

recent developments and reviews. Such action is necessary to allow for consistency in use and decision-making, and to assist regions which do not have the technical expertise or resources for such open-ended decision-making. Information provided in the letter of transmittal to the draft Manual only addresses this concern in part and is physically separated from the manual.

3.6 Tier IV Requirements

Presently, this tier is very poorly defined as regards both implementation and rationale/usage. Clearly "chronic" tests must include a substantial portion (if not all) of the life-cycle of a test organism. Endpoints must be related to population dynamics: survival, growth and reproduction are key parameters to measure at the individual organism level. Measures of growth must distinguish between somatic and gametic growth. This tier should determine and differentiate between statistical and ecological significance. Specifically, are the results of sensitive tests in tier III which indicate the potential for a concern likely to be realized in the environment? As such, there is a requirement for directed testing, not simply more testing. **The Subcommittee recommends that the manual describe some of the types of analyses and bioassays which may be included in Tier IV and discuss the rationale for such testing.**

4.0 SUMMARY OF RECOMMENDATIONS

a. The Subcommittee recommends that EPA revise its tiered approach to allow reevaluation at higher tiers of failures from lower tiers, to clarify the linkages between tiers, and to emphasize reducing uncertainty as the level of tier testing increases.

b. The Subcommittee recommends that QA/QC be a separate, detailed Section, and that protocols be appropriate and clearly specified.

c. The Subcommittee recommends that the manual describe some of the types of analyses and bioassays which may be included in Tier IV and discuss the rationale for such testing.

d. The Subcommittee recommends that a Glossary of Terms be generated and that the terms, as defined, be used in a consistent manner throughout the document.

e. The Subcommittee recommends elimination of redundancies and repetition such that important information is clear and not obscured.

f. The Subcommittee recommends that Table 11-1 be amended so that at least one species with an asterisk must be tested and adult bivalves should be removed from this list.

g. The Subcommittee recommends that Table 12-1 be amended such that recommended species be those in intimate contact with the sediments or which ingest sediments.

h. The Subcommittee questions whether the effort and emphasis on the *Neanthes arenaceodentata* "chronic" test is warranted. We do not question a general emphasis on polychaetes and amphipods for test development, as long as the tests are appropriately sensitive and useful.

i. The Subcommittee recommends that the inclusion of a sensitive infaunal amphipod or preferably an appropriately sensitive test incorporating measures of chronic sublethal effects, as stated in the Ocean Dumping Regulations, be required in Tier III. The tests should also be regionally appropriate.

j. The Subcommittee recommends that the issue of application factors be revisited and that EPA produce a peer-reviewed report that provides direction on this issue appropriate for the 1990s. Further we recommend that the manual clearly indicate that the 0.01 application factor is only potentially applicable to LC50s with persistent chemicals.

k. It is recommended that appropriate consensus guidelines be drafted for interpretation of bioaccumulation tests and that these guidelines be incorporated into the regulatory framework. It is further recommended that the Agency conduct studies that relate body burdens to "no observed effect levels" (NOEC) as part of chronic toxicity protocols.

l. The Subcommittee recommends that the estimation of the equilibrium concentrations for measuring bioaccumulation rates be based upon pharmacokinetic modeling.

m. The Subcommittee recommends that the Ocean Disposal Manual prohibit the use of reference sediments from the proposed disposal area, unless the area can be shown to be free of contamination. It is recommended that the reference site and sample selection guidance provided in pages 1-6 and 3-2 be expanded. Additional guidance is needed for determining when the "reference-point approach" should be used as compared to the "reference area approach".

n. The Subcommittee recommends that interpretative guidance, in the form of a framework for action, be provided now. This guidance should also provide the user with a list of critical changes that may be appropriate relative to recent developments and reviews.

o. Since it is not clear how SQC will be utilized in the disposal of dredged materials, we recommend that the applicability of sediment quality criteria to the new "Green Book" be defined.

5.0 LITERATURE CITED

Hartung, R. (1988). Assumptions and methods for assessing mixtures. Trace Subst. in Environmental Health. 22:194-204.

National Academy of Sciences (1972). "Water Quality Criteria 1972. A report of the Committee on Water Quality Criteria." National Academy of Sciences, Washington, D.C. EPA-R3-73-033. 594 pp. + 3 appendices.

National Research Council (1988). Complex Mixtures. Committee on Methods for *In Vivo* Toxicity Testing of Complex Mixtures. Board on Environmental Studies and Toxicology. National Academy Press xi + 227.

Stara, J.F. and L.S. Erdreich, editors (1984). Approaches to Risk Assessment for Multiple Chemical Exposures. Environ. Criteria & Assessment Office. EPA-600/9-84-008. viii + 277.

Stara, J.F. and L.S. Erdreich, editors (1985). Advances in Health Risk Assessment for Systemic Toxicants and Chemical Mixtures. Toxicol. Indust. Hlth 1(4):1-364.

U.S. EPA (1987). Guidelines for the health risk assessment of chemical mixtures. The Risk Assessment Guidelines of 1986. Office of Health & Environmental Assessment. EPA/600/8-87/045. pp. 3-1 to 3-3-16. (Also published under 51 FR 34014).

DISTRIBUTION LIST

The Administrator

Deputy Administrator

AA Office of Water
    Director, Office of Science and Technology
    Director, Office of Watersheds, Oceans, and Wetlands

AA Office of Research and Development
    Director, Office of Environmental Processes and Effects Research

Regional Administrators
EPA Headquarters Library
    EPA Regional Libraries