APPENDIX H

Derivation of the 1985 Aquatic Life Criteria

WATER QUALITY STANDARDS HANDBOOK

SECOND EDITION

Derivation of the 1985 Aquatic Life Critera

The following is a summary of the Guidelines for Derivation of Criteria for Aquatic Life. The complete text is found in "Guidelines for Deriving Numerical National Water Quality Criteria for the Protection of Aquatic Organisms and Their Uses," available from National Technical Information Service - PB85-227049.

Derivation of numerical national water quality criteria for the protection of aquatic organisms and their uses is a complex process that uses information from many areas of aquatic toxicology. When a national criterion is needed for a particular material, all available information concerning toxicity to and bioaccumulation by aquatic organisms is collected, reviewed for acceptability, and sorted. If enough acceptable data on acute toxicity to aquatic animals are available, they are used to estimate the highest one-hour average concentration that should not result in unacceptable effects on aquatic organisms and their uses. If justified, this concentration is made a function of water quality characteristics such as pH, salinity, or hardness. Similarly, data on the chronic toxicity of the material to aquatic animals are used to estimate the highest four-day average concentration that should not cause unacceptable toxicity during a long-term exposure. If appropriate, this concentration is also related to a water quality characteristic.

Data on toxicity to aquatic plants are examined to determine whether plants are likely to be unacceptably affected by concentrations that should not cause unacceptable effects on animals. Data on bioaccumulation by aquatic organisms are used to determine if residues might subject edible species to restrictions by the U.S. Food and Drug Administration (FDA), or if such residues might harm wildlife that consumes aquatic life. All other available data are examined for adverse effects that might be biologically important.

If a thorough review of the pertinent information indicates that enough acceptable data exists, numerical national water quality criteria are derived for fresh water or salt water or both to protect aquatic organisms and their uses from unacceptable effects due to exposures to high concentrations for short periods of time, lower concentrations for longer periods of time, and combinations of the two.

I. Definition of Material of Concern

A. Each separate chemical that does not ionize substantially in most natural bodies of water should usually be considered a separate material, except possibly for structurally similar organic compounds that exist only in large quantities as commercial mixtures of the various compounds and apparently have similar biological, chemical, physical, and toxicological properties.

B. For chemicals that do ionize substantially in most natural waterbodies (e.g., some phenols and organic acids, some salts of phenols and organic acids, and most inorganic salts and coordination complexes of metals), all forms in chemical equilibrium should usually be considered one material. Each different oxidation state of a metal and each different non-ionizable covalently bonded organometallic compound should usually be considered a separate material.

C. The definition of the material should include an operational analytical component. Identification of a material simply, for example, as "sodium" obviously implies "total sodium" but leaves room for doubt. If "total" is meant, it should be explicitly stated. Even

"total" has different operational definitions, some of which do not necessarily measure "all that is there" in all sample. Thus, it is also necessary to reference or describe one analytical method that is intended. The operational analytical component should take into account the analytical and environmental chemistry of the material, the desirability of using the same analytical method on samples from laboratory tests, ambient water and aqueous effluents, and various practical considerations such as labor and equipment requirements and whether the method would require measurement in the field or would allow measurement after samples are transported to a laboratory.

The primary requirements of the operational analytical component are that it be appropriate for use on samples of receiving water, compatible with the available toxicity and bioaccumulation data without making overly hypothetical extrapolations, and rarely result in underprotection or overprotection of aquatic organisms and their uses. Because an ideal analytical measurement will rarely be available, a compromise measurement will usually be used. This compromise measurement must fit with the general approach: if an ambient concentration is lower than the national criterion, unacceptable effects will probably not occur (i.e., the compromise measurement must not err on the side of underprotection when measurements are made on a surface water). Because the chemical and physical properties of an effluent are usually quite different from those of the receiving water, an analytical method acceptable for analyzing an effluent might not be appropriate for analyzing a receiving water, and vice versa. If the ambient concentration *calculated* from a measured concentration in an effluent is higher than the national criterion, an additional option is to *measure* the concentration after dilution of the effluent with receiving water to determine if the measured concentration is lowered by such phenomena as complexation or sorption. A further option, of course, is to derive a site-specific criterion (1,2,3). Thus, the criterion should be based on an appropriate analytical measurement, but the criterion is not rendered useless if an ideal measurement either is not available or is not feasible.

The analytical chemistry of the material might need to be considered when defining the material or when judging the acceptability of some toxicity tests, but a criterion should not be based on the sensitivity of an analytical method. When aquatic organisms are more sensitive than routine analytical methods, the proper solution is to develop better analytical methods, not to underprotect aquatic life.

II. Collection of Data

A. Collect all available data on the material concerning toxicity to, and bioaccumulation by, aquatic animals and plants; FDA action levels (compliance Policy Guide, U.S. Food & Drug Admin. 1981) and chronic feeding studies and long-term field studies with wildlife species that regularly consume aquatic organisms.

B. All data that are used should be available in typed, dated, and signed hard copy (publication, manuscript, letter, memorandum) with enough supporting information to indicate that acceptable test procedures were used and that the results are probably reliable. In some cases, additional written information from the investigator may be needed. Information that is confidential, privileged, or otherwise not available for distribution should not be used.

C. Questionable data, whether published or unpublished, should not be used. Examples would be data from tests that did not contain a control treatment, tests in which too many organisms in the control treatment died or showed signs of stress or disease, and tests in which distilled or deionized water was used as the dilution water without addition of appropriate salts.

D. Data on technical grade materials may be used, if appropriate; but data on formulated mixtures and emulsifiable concentrates of the material may not be used.

E. For some highly volatile, hydrolyzable, or degradable materials, only use data from flow-through tests in which the concentrations of test material were measured often enough with acceptable analytical methods.

F. Data should be rejected if obtained by using:

- Brine shrimp — because they usually occur naturally only in water with salinity greater than 35 g/kg;
- Species that do not have reproducing wild populations in North America; or
- Organisms that were previously exposed to substantial concentrations of the test material or other contaminants.

G. Questionable data, data on formulated mixtures and emulsifiable concentrates, and data obtained with nonresident species or previously exposed organisms may be used to provide auxiliary information but should not be used in the derivation of criteria.

III. Required Data

A. Certain data should be available to help ensure that each of the four major kinds of possible adverse effects receives adequate consideration: results of acute and chronic toxicity tests with representative species of aquatic animals are necessary to indicate the sensitivities of appropriate untested species. However, since procedures for conducting tests with aquatic plants and interpreting the results are not as well developed, fewer data concerning toxicity are required. Finally, data concerning bioaccumulation by aquatic organisms are required only with relevant information on the significance of residues in aquatic organisms.

B. To derive a criterion for freshwater aquatic organisms and their uses, the following should be available:

1. Results of acceptable acute tests (see section IV) with at least one species of freshwater animal in at least eight different families including all of the following:
   - The family Salmonidae in the class Osteichthyes.
   - A second family in the class Osteichthyes, preferably a commercially or recreationally important warmwater species, such as bluegill or channel catfish.
   - A third family in the phylum Chordata (may be in the class Osteichthyes or may be an amphibian, etc.).
   - A planktonic crustacean such as a cladoceran or copepod.
   - A benthic crustacean (ostracod, isopod, amphipod, crayfish, etc.).
   - An insect (mayfly, dragonfly, damselfly, stonefly, caddisfly, mosquito, midge, etc.).
   - A family in a phylum other than Arthropoda or Chordata, such as Rotifera, Annelida, Mollusca.
   - A family in any order of insect or any phylum not already represented.
2. Acute-chronic ratios (see section VI) with species of aquatic animals in at least three different families, provided that:
   - At least one is a fish;
   - At least one is an invertebrate; and
   - At least one is an acutely sensitive freshwater species (the other two may be saltwater species).
3. Results of at least one acceptable test with a freshwater alga or vascular plant (see section VIII). If the plants are among the aquatic organisms that are most sensitive to the material, test data on a plant in another phylum (division) should also be available.

4. At least one acceptable bioconcentration factor determined with an appropriate freshwater species, if a maximum permissible tissue concentration is available (see section IX).

C. To derive a criterion for saltwater aquatic organisms and their uses, the following should be available:

1. Results of acceptable acute tests (see section IV) with at least one species of saltwater animal in at least eight different families, including all of the following:
   - Two families in the phylum Chordata;
   - A family in a phylum other than Arthropoda or Chordata;
   - Either the Mysidae or Penaeidae family;
   - Three other families not in the phylum Chordata (may include Mysidae or Penaeidae, whichever was not used previously); and
   - Any other family.
2. Acute-chronic ratios (see section VI) with species of aquatic animals in at least three different families, provided that of the three species:
   - At least one is a fish;
   - At least one is an invertebrate; and
   - At least one is an acutely sensitive saltwater species (the other may be an acutely sensitive freshwater species).
3. Results of at least one acceptable test with a saltwater alga or vascular plant (see section VIII). If plants are among the aquatic organisms most sensitive to the material, results of a test with a plant in another phylum (division) should also be available.
4. At least one acceptable bioconcentration factor determined with an appropriate saltwater species, if a maximum permissible tissue concentration is available (see section IX).

D. If all required data are available, a numerical criterion can usually be derived, except in special cases. For example, derivation of a criterion might not be possible if the available acute-chronic ratios vary by more than a factor of 10 with no apparent pattern. Also, if a criterion is to be related to a water quality characteristic T (see sections V and VII), more data will be necessary.

Similarly, if all required data are not available, a numerical criterion should not be derived except in special cases. For example, even if not enough acute and chronic data are available, it might be possible to derive a criterion if the available data clearly indicate that the Final Residue Value should be much lower than either the Final Chronic Value or the Final Plant Value.

E. Confidence in a criterion usually increases as the amount of available pertinent data increases. Thus, additional data are usually desirable.

IV. Final Acute Value

A. Appropriate measures of the acute (short-term) toxicity of the material to a variety of species of aquatic animals are used to calculate the Final Acute Value. The Final Acute Value is an estimate of the concentration of the material, corresponding to a cumulative probability of 0.05 in the acute toxicity values for genera used in acceptable acute tests conducted on the material. However, in some cases, if the Species Mean Acute Value of a commercially or recreationally important species is lower than the calculated Final Acute Value, then that Species Mean Acute Value replaces the calculated Final Acute Value to protect that important species.

B. Acute toxicity tests should have been conducted using acceptable procedures (ASTM Standards E 729 and 724).

C. Except for tests with saltwater annelids and mysids, do not use results of acute tests during which test organisms were fed, unless data indicate that the food did not affect the toxicity of the test material.

D. Results of acute tests conducted in unusual dilution water (dilution water in which total organic carbon or particulate matter exceeded 5 mg/L) should not be used unless a relationship is developed between acute toxicity and organic carbon or particulate matter or unless data show that the organic carbon or particulate matter does not affect toxicity.

E. Acute values should be based on endpoints that reflect the total severe acute adverse impact of the test material on the organisms used in the test. Therefore, only the following kinds of data on acute toxicity to aquatic animals should be used:

1. Tests with daphnids and other cladocerans should be started with organisms less than 24-hours old, and tests with midges should be stressed with second- or third-instar larvae. The result should be the 48-hour EC_{50} based on percentage of organisms immobilized plus percentage of organisms killed. If such an EC_{50} is not available from a test, the 48-hour LC_{50} should be used in place of the desired 48-hour EC_{50}. An EC_{50} or LC_{50} of longer than 48 hours can be used as long as the animals were not fed and the control animals were acceptable at the end of the test.

2. The result of a test with embryos and larvae of barnacles, bivalve molluscs (clams, mussels, oysters, and scallops), sea urchins, lobsters, crabs, shrimp, and abalones should be the 96-hour EC_{50} based on the percentage of organisms with incompletely developed shells plus the percentage of organisms killed. If such an EC_{50} is not available from a test, the lower of the 96-hour EC_{50}, based on the percentage of organisms with incompletely developed shells and the 96-hour LC_{50} should be used in place of the desired 96-hour EC_{50}. If the duration of the test was between 48 and 96 hours, the EC_{50} or LC_{50} at the end of the test should be used.

3. The acute values from tests with all other freshwater and saltwater animal species and older life stages of barnacles, bivalve molluscs, sea urchins, lobsters, crabs, shrimps, and abalones should be the 96-hour EC_{50} based on the percentage of organisms exhibiting loss of equilibrium, plus the percentage of organisms immobilized, plus the percentage of organisms killed. If such an EC_{50} is not available from a test, the 96-hour LC_{50} should be used in place of the desired 96-hour EC_{50}.

4. Tests with single-celled organisms are not considered acute tests, even if the duration was 96 hours or less.

5. If the tests were conducted properly, acute values reported as "greater than" values and those above the solubility of the test material should be used because rejection of such acute values would unnecessarily lower the Final Acute Value by eliminating acute values for resistant species.

F. If the acute toxicity of the material to aquatic animals apparently has been shown to be related to a water quality characteristic such as hardness or particulate matter for freshwater animals or salinity or particulate matter for saltwater animals, a Final Acute Equation should be derived based on that water quality characteristic. (Go to section V.)

G. If the available data indicate that one or more life stages are at least a factor of 2 more resistant than one or more other life stages of the same species, the data for the more resistant life stages should not be used in the calculation of the Species Mean Acute Value because a species can be considered protected from acute toxicity only if all life stages are protected.

H. The agreement of the data within and between species should be considered. Acute values that appear to be questionable in comparison with other acute and chronic data for the same species and for other species in the same genus probably should not be used in

calculation of a Species Mean Acute Value. For example, if the acute values available for a species or genus differ by more than a factor of 10, some or all of the values probably should not be used in calculations.

I. For each species for which at least one acute value is available, the Species Mean Acute Value should be calculated as the geometric mean of the results of all flow-through tests in which the concentrations of test material were measured. For a species for which no such result is available, the Species Mean Acute Value should be calculated as the geometric mean of all available acute values — i.e., results of flow-through tests in which the concentrations were not measured and results of static and renewal tests based on initial concentrations of test material. (Nominal concentrations are acceptable for most test materials if measured concentrations are not available.)

NOTE: Data reported by original investigators should not be rounded off. Results of all intermediate calculations should be rounded to four significant digits.

NOTE: The geometric mean of N numbers is the N^{th} root of the product of the N numbers. Alternatively, the geometric mean can be calculated by adding the logarithms of the N numbers, dividing the sum by N, and taking the antilog of the quotient. The geometric mean of two numbers is the square root of the product of the two numbers, and the geometric mean of one number is that number. Either natural (base 0) or common (base 10) logarithms can be used to calculate geometric means as long as they are used consistently within each set of data (i.e., the antilog used must match the logarithm used).

NOTE: Geometric means rather than arithmetic means are used here because the distributions of individual organisms' sensitivities in toxicity tests on most materials, and the distributions of species' sensitivities within a genus, are more likely to be lognormal than normal. Similarly, geometric means are used for acute-chronic ratios and bioconcentration factors because quotients are likely to be closer to lognormal than normal distributions. In addition, division of the geometric mean of a set of numerators by the geometric mean of the set of corresponding denominators will result in the geometric mean of the set of corresponding quotients.

J. The Genus Mean Acute Value should be calculated as the geometric mean of the Species Mean Acute Values available for each genus.

K. Order the Genus Mean Acute Value from high to low.

L. Assign ranks, R, to the Genus Mean Acute Value from "1" for the lowest to "N" for the highest. If two or more Genus Mean Acute Values are identical, arbitrarily assign them successive ranks.

M. Calculate the cumulative probability, P, for each Genus Mean Acute Value as R/(N+1).

N. Select the four Genus Mean Acute Values that have cumulative probabilities closest to 0.05. (If there are less than 59 Genus Mean Acute Values, these will always be the four lowest Genus Mean Acute Values).

O. Using the selected Genus Mean Acute Values and Ps, calculate:

$$S^2 = \frac{\Sigma((\ln GMAV)^2) - ((\Sigma(\ln GMAV))^2/4)}{\Sigma(P) - ((\Sigma(\sqrt{P}))^2/4)}$$

$$L = (\Sigma(\ln GMAV) - S(\Sigma(\sqrt{P})))/4$$

$$A = S(\sqrt{0.05}) + L$$

$$FAV = e^A$$

(See original document, referenced at beginning of this appendix, for development of the calculation procedure and Appendix 2 for example calculation and computer program.)

NOTE: Natural logarithms (logarithms to base e, denoted as ln) are used herein merely because they are easier to use on some hand calculators and computers than common (base 10) logarithms. Consistent use of either will produce the same result.

P. If for a commercially or recreationally important species the geometric mean of the acute values from flow-through tests in which the concentrations of test material were measured is lower than the calculated Final Acute Value, then that geometric mean should be used as the Final Acute Value instead of the calculated Final Acute Value.

Q. Go to section VI.

V. Final Acute Equation

A. When enough data are available to show that acute toxicity to two or more species is similarly related to a water quality characteristic, the relationship should be taken into account as described in section IV, steps B through G, or using analysis of covariance. The two methods are equivalent and produce identical results. The manual method described below provides an understanding of this application of covariance analysis, but computerized versions of covariance analysis are much more convenient for analyzing large data tests. If two or more factors affect toxicity, multiple regression analysis should be used.

B. For each species for which comparable acute toxicity values are available at two or more different values of the water quality characteristic, perform a least squares regression of the acute toxicity values on the corresponding values of the water quality characteristic to obtain the slope and its 95 percent confidence limits for each species.

> NOTE: Because the best documented relationship fitting these data is that between hardness and acute toxicity of metals in freshwater and a log-log relationship, geometric means and natural logarithms of both toxicity and water quality are used in the rest of this section. For relationships based on other water quality characteristics such as pH, temperature, or salinity, no transformation or a different transformation might fit the data better, and appropriate changes will be necessary.

C. Decide whether the data for each species are useful, taking into account the range and number of the tested values of the water quality characteristic and the degree of agreement within and between species. For example, a slope based on six data points might be of limited value if based only on data for a very narrow range of water quality characteristic values. A slope based on only two data points, however, might be useful if consistent with other information and if the two points cover a broad enough range of the water quality characteristic.

In addition, acute values that appear to be questionable in comparison with other acute and chronic data available for the same species and for other species in the same genus probably should not be used. For example, if after adjustment for the water quality characteristic the acute values available for a species or genus differ by more than a factor of 10, probably some or all of the values should be rejected. If useful slopes are not available for at least one fish and one invertebrate, or if the available slopes are too dissimilar, or if too few data are available to adequately define the relationship between acute toxicity and the water quality characteristic, return to section IV.G, using the results of tests conducted under conditions and in waters similar to those commonly used for toxicity tests with the species.

D. Individually for each species, calculate the geometric mean of the available acute values and then divide each of these acute values by the mean for the species. This normalizes the values so that the geometric mean of the normalized values for each species, individually, and for any combination of species is 1.0.

E. Similarly normalize the values of the water quality characteristic for each species, individually.

F. Individually for each species, perform a least squares regression of the normalized acute toxicity values on the corresponding normalized values of the water quality characteristic. The resulting slopes and 95 percent confidence limits will be identical to those obtained in

step B. However, now, if the data are actually plotted, the line of best fit for each individual species will go through the point 1,1 in the center of the graph.

G. Treat normalized data as if they were all for the same species and perform a least squares regression of all the normalized acute values on the corresponding normalized values of the water quality characteristic to obtain the pooled acute slope, V, and its 95 percent confidence limits. If all the normalized data are actually plotted, the line of best fit will go through the point 1,1 in the center of the graph.

H. For each species, calculate the geometric mean, W, of the acute toxicity values and the geometric mean, X, of the values of the water quality characteristic. (These were calculated in steps D and E.)

I. For each species, calculate the logarithm, Y, of the Species Mean Acute Value at a selected value, Z, of the water quality characteristic using the equation:

$$Y = \ln W - V(\ln X - \ln Z).$$

J. For each species, calculate the SMAV at Z using the equation:

$$SMAV = e^{Y}.$$

NOTE: Alternatively, the Species Mean Acute Values at Z can be obtained by skipping step H using the equations in steps I and J to adjust each acute value individually to Z, and then calculating the geometric mean of the adjusted values for each species individually.

This alternative procedure allows an examination of the range of the adjusted acute values for each species.

K. Obtain the Final Acute Value at Z by using the procedure described in section IV, steps J through O.

L. If the Species Mean Acute Value at Z of a commercially or recreationally important species is lower than the calculated Final Acute Value at Z, then that Species Mean Acute Value should be used as the Final Acute Value at Z instead of the calculated Final Acute Value.

M. The Final Acute Equation is written as:

$$\text{Final Acute Value} = e^{(V[\ln(\text{water quality characteristic})] + \ln A - V[\ln Z])}$$

where

V = pooled acute slope

A = Final Acute Value at Z.

Because V, A, and Z are known, the Final Acute Value can be calculated for any selected value of the water quality characteristic.

VI. Final Chronic Value

A. Depending on the data that are available concerning chronic toxicity to aquatic animals, the Final Chronic Value might be calculated in the same manner as the Final Acute Value or by dividing the Final Acute Value by the Final Acute-Chronic Ratio. In some cases, it may not be possible to calculate a Final Chronic Value.

NOTE: As the name implies, the Acute-Chronic Ratio is a way of relating acute and chronic toxicities. The Acute-Chronic Ratio is basically the inverse of the application factor, but this new name is better because it is more descriptive and should help prevent confusion between "application factors" and "safety factors." Acute-Chronic Ratios and application factors are ways of relating the acute and chronic toxicities of a material to aquatic organisms. Safety factors are used to provide an extra margin of safety beyond the known or estimated sensitivities of aquatic organisms. Another advantage of the Acute-Chronic Ratio is that it will usually be greater than 1; this should avoid the confusion as to whether a large application factor is one that is close to unity or one that has a denominator that is much greater than the numerator.

B. Chronic values should be based on results of flow- through chronic tests in which the concentrations of test material in the test solutions were properly measured at appropriate times during the test. (Exception: renewal, which is acceptable for daphnids.)

C. Results of chronic tests in which survival, growth, or reproduction in the control treatment was unacceptably low should not be used. The limits of acceptability will depend on the species.

D. Results of chronic tests conducted in unusual dilution water (dilution water in which total organic carbon or particulate matter exceeded 5 mg/L) should not be used, unless a relationship is developed between chronic toxicity and organic carbon or particulate matter, or unless data show that organic carbon, particulate matter (and so forth) do not affect toxicity.

E. Chronic values should be based on endpoints and lengths of exposure appropriate to the species. Therefore, only results of the following kinds of chronic toxicity tests should be used:

1. Life-cycle toxicity tests consisting of exposures of each of two or more groups of individuals of a species to a different concentration of the test material throughout a life cycle. To ensure that all life stages and life processes are exposed, tests with fish should begin with embryos or newly hatched young less than 48-hours old, continue through maturation and reproduction, and end not less than 24 days (90 days for salmonids) after the hatching of the next generation. Tests with daphnids should begin with young less than 24-hours old and last for not less than 21 days. Tests with mysids should begin with young less than 24-hours old and continue until seven days past the median time of first brood release in the controls.

   For fish, data should be obtained and analyzed on survival and growth of adults and young, maturation of males and females, eggs spawned per female, embryo viability (salmonids only), and hatchability. For daphnids, data should be obtained and analyzed on survival and young per female. For mysids, data should be obtained and analyzed on survival, growth, and young per female.

2. Partial life-cycle toxicity tests consisting of exposures of each of two or more groups of individuals in a fish species to a concentration of the test material through most portions of a life cycle. Partial life-cycle tests are allowed with fish species that require more than a year to reach sexual maturity so that all major life stages can be exposed to the test material in less than 15 months.

   Exposure to the test material should begin with immature juveniles at least two months prior to active gonad development, continue through maturation and reproduction, and end not less than 24 days (90 days for salmonids) after the hatching of the next generation. Data should be obtained and analyzed on survival and growth of adults and young, maturation of males and females, eggs spawned per female, embryo viability (salmonids only), and hatchability.

3. Early life stage toxicity tests consisting of 28- to 32-day (60 days post hatch for salmonids) exposures of the early life stages of a fish species from shortly after fertilization through embryonic, larval, and early juvenile development. Data should be obtained and analyzed on survival and growth.

   NOTE: Results of an early life stage test are used as predictions of results of life-cycle and partial life-cycle tests with the same species. Therefore, when results of a total or partial life-cycle test are available, results of an early life stage test with the same species should not be used. Also, results of early life stage tests in which the incidence of mortalities or abnormalities increased substantially near the end should not be used because these results are possibly not good predictions of the results of comparable total or partial life cycle or partial life cycle tests.

F. A chronic value can be obtained by calculating the geometric mean of the lower and upper chronic limits from a chronic test or by analyzing chronic data using regression analysis. A lower chronic limit is the highest tested concentration in an acceptable chronic test that did not cause an unacceptable amount of adverse effect on any of the specified biological measurements and below which no tested concentration caused an unacceptable effect. An upper chronic limit is the lowest tested concentration in an acceptable chronic test that did cause an unacceptable amount of adverse effect on one or more of the specified biological measurements and above which all tested concentrations also caused such an effect.

NOTE: Because various authors have used a variety of terms and definitions to interpret and report results of chronic tests, reported results should be reviewed carefully. The amount of effect that is considered unacceptable is often based on a statistical hypothesis test but might also be defined in terms of a specified percent reduction from the controls. A small percent reduction (e.g., 3 percent) might be considered acceptable even if it is statistically significantly different from the control, whereas a large percent reduction (e.g., 30 percent) might be considered unacceptable even if it is not statistically significant.

G. If the chronic toxicity of the material to aquatic animals apparently has been shown to be related to a water quality characteristic such as hardness or particulate matter for freshwater animals or salinity or particulate matter for saltwater animals, a Final Chronic Equation should be derived based on that water quality characteristic. Go to section VII.

H. If chronic values are available for species in eight families as described in sections III.B.1 or III.C.1, a Species Mean Chronic Value should also be calculated for each species for which at least one chronic value is available by calculating the geometric mean of all chronic values available for the species; appropriate Genus Mean Chronic Values should also be calculated. The Final Chronic Value should then be obtained using the procedure described in section III, steps J through O. Then go to section VI.M.

I. For each chronic value for which at least one corresponding appropriate acute value is available, calculate an acute-chronic ratio using for the numerator the geometric mean of the results of all acceptable flow-through acute tests in the same dilution water and in which the concentrations were measured. (Exception: static is acceptable for daphnids.)

For fish, the acute test(s) should have been conducted with juveniles and should have been part of the same study as the chronic test. If acute tests were not conducted as part of the same study, acute tests conducted in the same laboratory and dilution water but in a different study may be used. If no such acute tests are available, results of acute tests conducted in the same dilution water in a different laboratory may be used. If no such acute tests are available, an acute-chronic ratio should not be calculated.

J. For each species, calculate the species mean acute-chronic ratio as the geometric mean of all acute-chronic ratios available for that species.

K. For some materials, the acute-chronic ratio seems to be the same for all species, but for other materials, the ratio seems to increase or decrease as the Species Mean Acute Value increases. Thus the Final Acute-Chronic Ratio can be obtained in four ways, depending on the data available:

1. If the Species Mean Acute-Chronic ratio seems to increase or decrease as the Species Mean Acute Value increases, the Final Acute-Chronic Ratio should be calculated as the geometric mean of the acute-chronic ratios for species whose Species Mean Acute Values are close to the Final Acute Value.

2. If no major trend is apparent, and the acute-chronic ratios for a number of species are within a factor of 10, the Final Acute-Chronic Ratio should be calculated as the geometric mean of all the Species Mean Acute-Chronic Ratios available for both freshwater and saltwater species.

3. For acute tests conducted on metals and possibly other substances with embryos and larvae of barnacles, bivalve molluscs, sea urchins, lobsters, crabs, shrimp, and abalones (see section IV.E.2), it is probably appropriate to assume that the

acute-chronic ratio is 2. Chronic tests are very difficult to conduct with most such species, but the sensitivities of embryos and larvae would likely determine the results of life cycle tests. Thus, if the lowest available Species Mean Acute Values were determined with embryos and larvae of such species, the Final Acute-Chronic Ratio should probably be assumed to be 2, so that the Final Chronic Value is equal to the Criterion Maximum Concentration (see section XI.B)

4. If the most appropriate Species Mean Acute-Chronic Ratios are less than 2.0, and especially if they are less than 1.0, acclimation has probably occurred during the chronic test. Because continuous exposure and acclimation cannot be assured to provide adequate protection in field situations, the Final Acute-Chronic Ratio should be assumed to be 2, so that the Final Chronic Value is equal to the Criterion Maximum Concentration (see section XI.B).

   If the available Species Mean Acute-Chronic Ratios do not fit one of these cases, a Final Acute-Chronic Ratio probably cannot be obtained, and a Final Chronic Value probably cannot be calculated.

L. Calculate the Final Chronic Value by dividing the Final Acute Value by the Final Acute-Chronic Ratio. If there was a Final Acute Equation rather than a Final Acute Value, see also section VII.A.

M. If the Species Mean Chronic Value of a commercially or recreationally important species is lower than the calculated Final Chronic Value, then that Species Mean Chronic Value should be used as the Final Chronic Value instead of the calculated Final Chronic Value.

N. Go to section VIII.

VII. Final Chronic Equation

A. A Final Chronic Equation can be derived in two ways. The procedure described here will result in the chronic slope being the same as the acute slope. The procedure described in steps B through N usually will result in the chronic slope being different from the acute slope.

1. If acute-chronic ratios are available for enough species at enough values of the water quality characteristic to indicate that the acute-chronic ratio is probably the same for all species and is probably independent of the water quality characteristic, calculate the Final Acute-Chronic Ratio as the geometric mean of the available Species Mean Acute-Chronic Ratios.

2. Calculate the Final Chronic Value at the selected value Z of the water quality characteristic by dividing the Final Acute Value at Z (see section V.M) by the Final Acute-Chronic Ratio.

3. Use V = pooled acute slope (see section V.M) as L = pooled chronic slope.

4. Go to section VII.M.

B. When enough data are available to show that chronic toxicity to at least one species is related to a water quality characteristic, the relationship should be taken into account as described in steps B through G or using analysis of covariance. The two methods are equivalent and produce identical results. The manual method described in the next paragraph provides an understanding of this application of covariance analysis, but computerized versions of covariance analysis are much more convenient for analyzing large data sets. If two or more factors affect toxicity, multiple regression analysis should be used.

C. For each species for which comparable chronic toxicity values are available at two or more different values of the water quality characteristic, perform a least squares regression of

the chronic toxicity values on the corresponding values of the water quality characteristic to obtain the slope and its 95 percent confidence limits for each species.

NOTE: Because the best-documented relationship fitting these data is that between hardness and acute toxicity of metals in fresh water and a log-log relationship, geometric means and natural logarithms of both toxicity and water quality are used in the rest of this section. For relationships based on other water quality characteristics such as pH, temperature, or salinity, no transformation or a different transformation might fit the data better, and appropriate changes will be necessary throughout this section. It is probably preferable, but not necessary, to use the same transformation that was used with the acute values in section V.

D. Decide whether the data for each species are useful, taking into account the range and number of the tested values of the water quality characteristic and the degree of agreement within and between species. For example, a slope based on six data points might be of limited value if founded only on data for a very narrow range of values of the water quality characteristic. A slope based on only two data points, however, might be useful if it is consistent with other information and if the two points cover a broad enough range of the water quality characteristic. In addition, chronic values that appear to be questionable in comparison with other acute and chronic data available for the same species and for other species in the same genus probably should not be used. For example, if after adjustment for the water quality characteristic the chronic values available for a species or genus differ by more than a factor of 10, probably some or all of the values should be rejected.

If a useful chronic slope is not available for at least one species, or if the available slopes are too dissimilar, or if too few data are available to adequately define the relationship between chronic toxicity and the water quality characteristic, the chronic slope is probably the same as the acute slope, which is equivalent to assuming that the acute-chronic ratio is independent of the water quality characteristic. Alternatively, return to section VI.H, using the results of tests conducted under conditions and in waters similar to those commonly used for toxicity tests with the species.

E. Individually for each species, calculate the geometric mean of the available chronic values and then divide each chronic value for a species by its mean. This normalizes the chronic values so that the geometric mean of the normalized values for each species individually, and for any combination of species, is 1.0.

F. Similarly normalize the values of the water quality characteristic for each species, individually.

G. Individually for each species, perform a least squares regression of the normalized chronic toxicity values on the corresponding normalized values of the water quality characteristic. The resulting slopes and the 95 percent confidence limits will be identical to those obtained in section B. Now, however, if the data are actually plotted, the line of best fit for each individual species will go through the point 1,1 in the center of the graph.

H. Treat all the normalized data as if they were all for the same species and perform a least squares regression of all the normalized chronic values on the corresponding normalized values of the water quality characteristic to obtain the pooled chronic slope, L, and its 95 percent confidence limits. If all the normalized data are actually plotted, the line of best fit will go through the point 1,1 in the center of the graph.

I. For each species, calculate the geometric mean, M, of the toxicity values and the geometric mean, P, of the values of the water quality characteristic. (These were calculated in steps E and F.)

J. For each species, calculate the logarithm, Q, of the Species Mean Chronic Value at a selected value, Z, of the water quality characteristic using the equation:

$$Q = \ln M - L(\ln P - \ln Z).$$

NOTE: Although it is not necessary, it will usually be best to use the same value of the water quality characteristic here as was used in section V.I.

K. For each species, calculate a Species Mean Chronic Value at Z using the equation:

$$SMCV = e^{Q}.$$

NOTE: Alternatively, the Species Mean Chronic Values at Z can be obtained by skipping step J, using the equations in steps J and K to adjust each acute value individually to Z, and then calculating the geometric means of the adjusted values for each species individually. This alternative procedure allows an examination of the range of the adjusted chronic values for each species.

L. Obtain the Final Chronic Value at Z by using the procedure described in section IV, steps J through O.

M. If the Species Mean Chronic Value at Z of a commercially or recreationally important species is lower than the calculated Final Chronic Value at Z, then that Species Mean Chronic Value should be used as the Final Chronic Value at Z instead of the calculated Final Chronic Value.

N. The Final Chronic Equation is written as:

$$\text{Final Chronic Value} = e^{(L[\ln(\text{water quality characteristic})] + \ln S - L[\ln Z])}$$

where

L = pooled chronic slope

S = Final Chronic Value at Z.

Because L, S, and Z are known, the Final Chronic Value can be calculated for any selected value of the water quality characteristic.

VIII. Final Plant Value

A. Appropriate measures of the toxicity of the material to aquatic plants are used to compare the relative sensitivities of aquatic plants and animals. Although procedures for conducting and interpreting the results of toxicity tests with plants are not well developed, results of tests with plants usually indicate that criteria which adequately protect aquatic animals and their uses will probably also protect aquatic plants and their uses.

B. A plant value is the result of a 96-hour test conducted with an alga, or a chronic test conducted with an aquatic vascular plant.

NOTE: A test of the toxicity of a metal to a plant usually should not be used if the medium contained an excessive amount of a complexing agent, such as EDTA, that might affect the toxicity of the metal. Concentrations of EDTA above about 200 µg/L should probably be considered excessive.

C. The Final Plant Value should be obtained by selecting the lowest result from a test with an important aquatic plant species in which the concentrations of test material were measured, and the endpoint was biologically important.

IX. Final Residue Value

A. The Final Residue Value is intended to prevent concentrations in commercially or recreationally important aquatic species from affecting marketability because they exceed applicable FDA action levels and to protect wildlife (including fishes and birds) that consume aquatic organisms from demonstrated unacceptable effects. The Final Residue Value is the lowest of the residue values that are obtained by dividing maximum permissible tissue concentrations by appropriate bioconcentration or bioaccumulation factors. A maximum permissible tissue concentration is either (a) an FDA action level (Compliance Policy Guide, U.S. Food & Drug Admin. 1981) for fish oil or for the edible portion of fish or shellfish, or a maximum acceptable dietary intake based on observations on survival, growth, or reproduction in a chronic wildlife feeding study or a long-term wildlife field study. If no maximum permissible tissue concentration is available, go to section X because no Final Residue Value can be derived.

B. Bioconcentration Factors (BCFs) and bioaccumulation factors (BAFs) are quotients of the concentration of a material in one or more tissues of an aquatic organism, divided by the average concentration in the solution in which the organism had been living. A BCF is intended to account only for net uptake directly from water and thus almost must be measured in a laboratory test. Some uptake during the bioconcentration test might not be directly from water if the food sorbs some of the test material before it is eaten by the test organisms. A BAF is intended to account for net uptake from both food and water in a real-world situation. A BAF almost must be measured in a field situation in which predators accumulate the material directly from water and by consuming prey that could have accumulated the material from both food and water.

The BCF and BAF are probably similar for a material with a low BCF, but the BAF is probably higher than the BCF for materials with high BCFs. Although BCFs are not too difficult to determine, very few BAFs have been measured acceptably because adequate measurements must be made of the material's concentration in water to ascertain if it was reasonably constant for a long enough time over the range of territory inhabited by the organisms. Because so few acceptable BAFs are available, only BCFs will be discussed further. However, if an acceptable BAF is available for a material, it should be used instead of any available BCFs.

C. If a maximum permissible tissue concentration is available for a substance (e.g., parent material, parent material plus metabolites, etc.), the tissue concentration used in the calculation of the BCF should be for the same substance. Otherwise, the tissue concentration used in the calculation of the BCF should derive from the material and its metabolites that are structurally similar and are not much more soluble in water than the parent material.

1. A BCF should be used only if the test was flow-through, the BCF was calculated based on measured concentrations of the test material in tissue and in the test solution, and the exposure continued at least until either apparent steady state or 28 days was reached. Steady state is reached when the BCF does not change significantly over a period of time, such as 2 days or 16 percent of the length of the exposure, whichever is longer. The BCF used from a test should be the highest of the apparent steady-state BCF, if apparent steady state was reached; the highest BCF obtained, if apparent steady state was not reached; and the projected steady state BCF, if calculated.

2. Whenever a BCF is determined for a lipophilic material, the percent lipids should also be determined in the tissue(s) for which the BCF was calculated.

3. A BCF obtained from an exposure that adversely affected the test organisms may be used only if it is similar to a BCF obtained with unaffected organisms of the same species at lower concentrations that did not cause adverse effects.

4. Because maximum permissible tissue concentrations are almost never based on dry weights, a BCF calculated using dry tissue weights must be converted to a wet tissue weight basis. If no conversion factor is reported with the BCF, multiply the dry weight BCF by 0.1 for plankton and by 0.2 for individual species of fishes and invertebrates.

5. If more than one acceptable BCF is available for a species, the geometric mean of the available values should be used; however, the BCFs are from different lengths of exposure and the BCF increases with length of exposure, then the BCF for the longest exposure should be used.

E. If enough pertinent data exists, several residue values can be calculated by dividing maximum permissible tissue concentrations by appropriate BCFs:

1. For each available maximum acceptable dietary intake derived from a chronic feeding study or a long-term field study with wildlife (including birds and aquatic organisms), the appropriate BCF is based on the whole body of aquatic species that constitutes or represents a major portion of the diet of the tested wildlife species.

2. For an FDA action level for fish or shellfish, the appropriate BCF is the highest geometric mean species BCF for the edible portion (muscle for decapods, muscle with or without skin for fishes, adductor muscle for scallops, and total soft tissue for other bivalve molluscs) of a consumed species. The highest species BCF is used because FDA action levels are applied on a species-by-species basis.

F. For lipophilic materials, calculating additional residue values is possible. Because the steady-state BCF for a lipophilic material seems to be proportional to percent lipids from one tissue to another and from one species to another, extrapolations can be made from tested tissues, or species to untested tissues, or species on the basis of percent lipids.

1. For each BCF for which the percent lipids is known for the same tissue for which the BCF was measured, normalize the BCF to a 1 percent lipid basis by dividing it by the percent lipids. This adjustment to a 1 percent lipid basis is intended to make all the measured BCFs for a material comparable regardless of the species or tissue with which the BCF was measured.
2. Calculate the geometric mean-normalized BCF. Data for both saltwater and freshwater species should be used to determine the mean-normalized BCF unless they show that the normalized BCFs are probably not similar.
3. Calculate all possible residue values by dividing the available maximum permissible tissue concentrations by the mean-normalized BCF and by the percent lipids values appropriate to the maximum permissible tissue concentrations, i.e.,

$$\text{Residue value} = \frac{(\text{maximum permissible tissue concentration})}{(\text{mean normalized BCF})(\text{appropriate percent lipids})}$$

- For an FDA action level for fish oil, the appropriate percent lipids value is 100.
- For an FDA action level for fish, the appropriate percent lipids value is 11 for freshwater criteria and 10 for saltwater criteria because FDA action levels are applied species-by-species to commonly consumed species. The highest lipid contents in the edible portions of important consumed species are about 11 percent for both the freshwater chinook salmon and lake trout and about 10 percent for the saltwater Atlantic herring.
- For a maximum acceptable dietary intake derived from a chronic feeding study or a long-term field study with wildlife, the appropriate percent lipids is that of an aquatic species or group of aquatic species that constitute a major portion of the diet of the wildlife species.

G. The Final Residue Value is obtained by selecting the lowest of the available residue values.

NOTE: In some cases, the Final Residue Value will not be low enough. For example, a residue value calculated from a FDA action level will probably result in an average concentration in the edible portion of a fatty species at the action level. Some individual organisms and possibly some species will have residue concentrations higher than the mean value, but no mechanism has been devised to provide appropriate additional protection. Also, some chronic feeding studies and long-term field studies with wildlife identify concentrations that cause adverse effects but do not identify concentrations that do not cause adverse effects; again, no mechanism has been devised to provide appropriate additional protection. These are some of the species and uses that are not protected at all times in all places.

X. Other Data

Pertinent information that could not be used in earlier sections might be available concerning adverse effects on aquatic organisms and their uses. The most important of these are data on cumulative and delayed toxicity, flavor impairment, reduction in survival, growth, or reproduction, or any other adverse effect shown to be biologically important. Especially important are data for species for which no other data are available. Data from behavioral, biochemical, physiological, microcosm, and field studies might also be available. Data might be available from tests conducted in unusual dilution water (see IV.D and VI.D), from chronic tests

in which the concentrations were not measured (see VI.B), from tests with previously exposed organisms (see II.F), and from tests on formulated mixtures or emulsifiable concentrates (see II.D). Such data might affect a criterion if they were obtained with an important species, the test concentrations were measured, and the endpoint was biologically important.

XI. Criterion

A. A criterion consists of two concentrations: the Criterion Maximum Concentration and the Criterion Continuous Concentration.

B. The Criterion Maximum Concentration (CMC) is equal to one-half the Final Acute Value.

C. The Criterion Continuous Concentration (CCC) is equal to the lowest of the Final Chronic Value, the Final Plant Value, and the Final Residue Value, unless other data (see section X) show that a lower value should be used. If toxicity is related to a water quality characteristic, the Criterion Continuous Concentration is obtained from the Final Chronic Equation, the Final Plant Value, and the Final Residue Value by selecting the one, or the combination, that results in the lowest concentrations in the usual range of the water quality characteristic, unless other data (see section X) show that a lower value should be used.

D. Round both the Criterion Maximum Concentration and the Criterion Continuous Concentration to two significant digits.

E. The criterion is stated as follows:

The procedures described in the "Guidelines for Deriving Numerical National Water Quality Criteria for the Protection of Aquatic Organisms and Their Uses" indicate that, except possibly where a locally important species is very sensitive, (1) aquatic organisms and their uses should not be affected unacceptably if the four-day average concentration of (2) does not exceed (3) μg/L more than once every three years on the average, and if the one-hour average concentration does not exceed (4) μg/L more than once every three years on the average.

where (1) = insert freshwater or saltwater

(2) = insert name of material

(3) = insert the Criterion Continuous Concentration

(4) = insert the Criterion Maximum Concentration.

XII. Final Review

A. The derivation of the criterion should be carefully reviewed by rechecking each step of the guidelines. Items that should be especially checked are

1. If unpublished data are used, are they well documented?
2. Are all required data available?
3. Is the range of acute values for any species greater than a factor of 10?
4. Is the range of Species Mean Acute Values for any genus greater than a factor of 10?
5. Is there more than a factor of 10 difference between the four lowest Genus Mean Acute Values?
6. Are any of the four lowest Genus Mean Acute Values questionable?
7. Is the Final Acute Value reasonable in comparison with the Species Mean Acute Values and Genus Mean Acute Values?
8. For any commercially or recreationally important species, is the geometric mean of the acute values from flow-through tests in which the concentrations of test material were measured lower than the Final Acute Value?

9. Are any of the chronic values questionable?
10. Are chronic values available for acutely sensitive species?
11. Is the range of acute-chronic ratios greater than a factor of 10?
12. Is the Final Chronic Value reasonable in comparison with the available acute and chronic data?
13. Is the measured or predicted chronic value for any commercially or recreationally important species below the Final Chronic Value?
14. Are any of the other data important?
15. Do any data look like they might be outliers?
16. Are there any deviations from the guidelines? Are they acceptable?

B. On the basis of all available pertinent laboratory and field information, determine if the criterion is consistent with sound scientific evidence. If not, another criterion — either higher or lower — should be derived using appropriate modifications of these guidelines.