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

a. For an FDA action level for fish oil, the appropriate percent lipids value is 100.

b. For an FDA action level for fish, the appropriate percent lipids value is 15 for freshwater criteria and 16 for saltwater criteria because FDA action levels are applied on a species-by-species basis to commonly consumed species. The edible portion of the freshwater lake trout averages about 15 percent lipids, and the edible portion of the saltwater Atlantic herring averages about 16 percent lipids (Sidwell, V. D., et al. 1974 Composition of the Edible Portion of Raw (Fresh or Frozen) Crustaceans, Finfish, and Mollusks. I. Protein, Fat, Moisture, Ash, Carbohydrate, Energy Value, and Cholesterol. Marine Fisheries Review 36:21–35).

c. For a maximum acceptable dietary intake derived from a chronic feeding study with wildlife, the appropriate percent lipids is the percent lipids of an aquatic species or group of aquatic species which constitute a major portion of the diet of the wildlife species.

F. The Final Residue Value is obtained by selecting the lowest of the available residue values. It should be noted that in many cases the Final Residue Value will not be low enough. For example, a residue value calculated from an FDA action level would result in an average concentration in the edible portion of a fatty species that is at the action level. On the average half of the individuals of the species would have concentrations above the FDA action level. Also, the results of many chronic feeding studies are concentrations that cause adverse effects.

X. Other Data

Pertinent information that could not be used in earlier sections may be available concerning adverse effects on aquatic organisms and their uses. The most important of these are data on flavor impairment, reduction in survival, growth, or reproduction, or any other adverse effect that has been shown to be biologically significant. Especially important are data for species for which no other data are available. Data from behavioral, micorcosm, field, and physiological studies may also be available.

XI. Criterion

A. The criterion consists of two concentrations, one that should not be exceeded on the average in a 24-hour period and one that should not be exceeded at any time during the 24-hour period. This two-number criterion is intended to identify water quality conditions that should protect aquatic life and its uses from acute and chronic adverse effects of both cumulative and noncumulative substances without being as restrictive as a one-number criterion would have to be to provide the same degree of protection.

B. The maximum concentration is the Final Acute Value or is obtained from the Final Acute Equation.

C. The 24-hour average concentration is obtained from the Final Chronic Value, the Final Plant Value, and the Final Residue Value by selecting the lowest available value, unless other data (see Section X) from tests in which the toxicant concentrations were measured show that a lower value should be used. If toxicity is related to a water quality characteristic, the 24-hour average concentration is obtained from the Final Chronic Equation, the Final Plant Value, and the Final Residue Value by selecting the one that results in the lowest concentrations in the normal range of the water quality characteristic, unless other data (see Section X) from tests in which the toxicant concentrations were measured show that a lower value should be used.

D. The criterion is (the 24-hour average concentration) as a 24-hour average and the concentration should not exceed (the maximum concentration) at any time.

XII. Review

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

These Guidelines were written by Charles E. Stephan, Donald I. Mount, David J. Hansen, John H. Gentile, Gary A. Chapman and William A. Brungs of the U.S.E.P.A. Environmental Research Laboratories in Corvallis, Oregon, Duluth, Minnesota, Gulf Breeze, Florida, and Narragansett, Rhode Island. Numerous other people, many of whom do not work for U.S.E.P.A., provided assistance and suggestions.

Appendix C–Guidelines and Methodology Used in the Preparation of Health Effect Assessment Chapters of the Consent Decree Water Criteria Documents

I. Objective

The objective of the health effect assessment chapters of the ambient water criteria documents is to estimate ambient water concentrations which do not represent a significant risk to the public. These assessments should constitute a review of all relevant information on individual chemicals or chemical classes in order to derive criteria that represent, in the case of suspect or proven carcinogens, various levels of incremental cancer risk, or, in the case of other pollutants, estimates of no-effect levels.

Ideally, ambient water quality criteria should represent levels for compounds in ambient water that do not pose a hazard to the human population. However, in any realistic assessment of human health hazard, a fundamental distinction must be made between absolute safety and the recognition of some risk. Criteria for absolute safety would have to be based on detailed knowledge of dose-response relationships in humans, including all sources of chemical exposure, the types of toxic effects elicited, the existence of thresholds for the toxic effects, the significance of toxicant interactions, and the variances of sensitivities and exposure levels within the human population. In practice, such absolute criteria cannot be established because of deficiencies in both the available data and the means of interpreting this information. Consequently, the individual human health effects chapters propose criteria which minimize or specify the potential risk of adverse human effects due to substances in ambient water. Potential social or economic costs and benefits are not considered in the formulation of the criteria.

II. Types of Criteria

Ambient water quality criteria are based on three types of biological endpoints: carcinogenicity, toxicity (i.e., all adverse effects other than cancer), and organoleptic effects.

For the purpose of deriving ambient water quality criteria, carcinogenicity is regarded as a non-threshold phenomenon. Using this assumption, "safe" or "no effect" levels for carcinogens cannot be established because even extremely small doses must be assumed to elicit a finite increase in the incidence of the response. Consequently, water quality criteria for carcinogens are presented as a range of pollutant concentrations associated with corresponding incremental risks.

For compounds which do not manifest any apparent carcinogenic effect, the threshold assumption is used in deriving a criterion. This assumption is based on the premise that a physiological reserve capacity exists within the organism which is thought to be depleted before clinical disease ensues. Alternatively, it may be assumed that the rate of damage will be insignificant over the life span of the organism. Thus, ambient water quality criteria are derived for non-carcinogenic chemicals, and presumably result in no observable-adverse-effect levels (NOAELs) in the exposed human population.

In some instances, criteria are based on organoleptic characteristics, i.e., thresholds for taste or odor. Such criteria are established when insufficient information is available on toxicologic effects or when the estimate of the level of the pollutant in ambient water based on organoleptic effects is lower than the level calculated from toxicologic data. It should be recognized that criteria based solely on organoleptic effects do not necessarily represent approximations of acceptable risk levels for human health.

Several ambient water quality criteria documents deal with classes of compounds which include chemicals exhibiting varying degrees of structural similarity. Because prediction of biological effects based solely on structural parameters is difficult, the derivation of compound-specific criteria is preferable to a class criterion. A compound-specific criterion is defined as a level derived from data on each individual subject compound that does not represent a significant risk to the public. For some chemical classes, however, a compound-specific criterion cannot be derived for each member of a class. In such instances, it is sometimes justifiable to derive a class criterion in which available data on one member of a class may be used to estimate criteria for other chemicals of the class because a sufficient data base is not available for those compounds.

For some chemicals and chemical classes, the data base was judged to be insufficient for the derivation of a criterion. In those cases, deficiencies in the available information are detailed.

III. Approach

The human health effects chapters attempt to summarize all information on the individual chemicals or classes of chemicals which might be useful in the risk assessment process to develop water quality criteria. Although primary emphasis is placed on identifying epidemiologic and toxicologic data, these assessments typically contain discussions on four topics: existing levels of human exposure, pharmacokinetics, toxic effects, and criterion formulation.

For all documents, an attempt is made to include the known relevant information. Review articles and reports are often used in the process of data evaluation and synthesis. Scientific judgment is exercised in the review and evaluation of the data in each document and in the identification of the adverse effects against which protective criteria are sought. In addition, each of these documents is reviewed by a peer committee of scientists familiar with the specific compound(s). These work-groups evaluate the quality of the available data, the completeness of the data summary, and the validity of the derived criterion.

In the analysis and organization of the data, an attempt is made to be consistent with respect to the format and the application of acceptable scientific principles. Evaluation procedures used in the hazard assessment process follow the principles outlined by the National Academy of Sciences in *Drinking Water and Health* (1977) and the guidelines of the Carcinogen Assessment Group of the U.S. EPA.

A. Exposure

The exposure section of the health effects chapters reviews known information on current levels of human exposure to the individual pollutant from all sources. Much of the data was obtained from monitoring studies of air, water, food, soil, and human or animal tissue residues. The major purpose of this section is to provide background information on the contribution of water exposure relative to all other sources. Consequently, the exposure section includes subsections reviewing different routes of exposure including water and food ingestion, inhalation, and dermal contact.

Information on exposure can be valuable in developing and assessing a water quality criterion. In these documents exposure from consumption of contaminated water and contaminated fish and shellfish products is used in criterion formulation. Data for all modes of exposure are useful in relating total intake to the expected contribution from contaminated water, fish, and shellfish. In addition, information for all routes of exposure, not limited to drinking water and fish and shellfish ingestion, can be used to justify or assess the feasibility of the formulation of criteria for ambient water.

The use of fish consumption as an exposure factor requires the quantitation of pollutant residues in the edible portions of the ingested species. Accordingly, bioconcentration factors (BCFs) are used to relate pollutant residues in aquatic organisms to the pollutant concentration in the ambient waters in which they reside.

To estimate the average per capita intake of a pollutant due to consumption of contaminated fish and shellfish the results of a diet survey were analyzed to calculate the average consumption of freshwater and estuarine fish and shellfish (U.S. EPA, 1980). A species is considered to be a consumed freshwater or estuarine fish and shellfish species if at some stage in its life cycle, it is harvested from fresh or estuarine water for human consumption in significant quantities (Stephan, 1980).

Three different procedures are used to estimate the weighted average BCF depending upon the lipid solubility of the chemical and the availability of bioconcentration data.

For lipid-soluble compounds, the average BCF is calculated from the weighted average percent lipids in the edible portions of consumed freshwater and estuarine fish and shellfish which was calculated from data on consumption of each species and its corresponding percent lipids to be 3.0 percent (Stephan, 1980). Because the steady-state BCFs for lipid-soluble compounds are proportional to percent lipids, bioconcentration factors for fish and shellfish can be adjusted to the average percent lipids for aquatic organisms consumed by Americans. For many lipid-soluble pollutants, there exists at least one BCF for which the percent lipid value was measured for the tissues for which the BCF is determined.

With 3.0 percent as the weighted average percent lipids for freshwater and estuarine fish and shellfish in the average diet, a BCF, and a corresponding percent lipid value, the weighted average bioconcentration factor can be calculated.

Example:

Weighted average percent lipids for average diet=3.0 percent

Measured BCF of 17 for trichloroethylene with bluegills at 4.8 percent lipids

Weighted average BCF for average diet equals

$$17 \times \frac{3.0\%}{4.8\%} = 10.6$$

As an estimate, 10.6 is used for the BCF.

In those cases where an appropriate bioconcentration factor is not available, the equation "Log BCF=(0.85 Log P)-0.70" can be used (Veith, et al. 1979) to estimate the BCF for aquatic organisms containing about 7.6 percent lipids (Veith, 1980) from the octanol/water partition coefficient P. An adjustment for percent lipids in the average diet versus 7.6 percent is made in order to derive the weighted average bioconcentration factor.

For non-lipid-soluble compounds, the available BCFs for the edible portion of consumed freshwater and estuarine fish and shellfish are weighted according to consumption factors to determine a weighted BCF representative of the average diet.

B. Pharmacokinetics

This section summarizes the available information on the absorption, distribution, metabolism, and elimination of the compound(s) in humans and experimental mammals. Conceptually, such information is useful in validation of inter- and intraspecies extrapolations, and in characterizing the modes of toxic action. Sufficient information on absorption and excretion in animals, together with a knowledge of ambient concentrations in water, food, and air, could be useful in estimating body burdens of chemicals in the human population. Distribution data which suggest target organs or tissues are desirable for interspecies comparison techniques. In terms of the derivation of criteria, pharmacokinetic data are essential to estimate equivalent oral doses based on data from inhalation or other routes of exposure.

C. Effects

This section summarizes information on biological effects in both humans and experimental mammals resulting in: acute, subacute, and chronic toxicity, synergism and/or antagonism, teratogenicity, mutagenicity, or carcinogenicity.

The major goal of this section is to survey the suitability of the data for use in assessment of hazard and to determine which biological end-point, i.e., non-threshold, threshold, or organoleptic, should be selected for use in criterion formulation.

Because this section attempts to assess potential human health effects, data on documented human effects are thoroughly evaluated. However, several factors inherent in human epidemiological studies usually preclude the use of such data in generating water quality criteria. These problems, as summarized by the National Academy of Sciences (NAS, 1977) are as follows:

1. Epidemiology cannot tell what effects a material will have until after humans have been exposed. One must not conduct what might be hazardous experiments on man.
2. If exposure has been ubiquitous, it may be impossible to assess the effects of a material, because there is no unexposed control group. Statistics of morbidity obtained before use of a new material can sometimes be useful, but when latent periods are variable and times of introduction and removal of materials overlap, historical data on chronic effects are usually unsatisfactory.
3. It is usually difficult to determine doses in human exposures.
4. Usually, it is hard to identify small changes in common effects, which may nonetheless be important if the population is large.
5. Interactions in a "nature-designed" experiment usually cannot be controlled.

Although these problems often prevent the use of epidemiological data in quantitative risk assessments, qualitative similarities or differences between documented effects in humans and observed effects in experimental mammals are extremely useful in testing the validity of animal-to-man extrapolations. Consequently, in each case, an attempt is made to identify and utilize both epidemiologic and animal dose-response data. Criteria derived from such a confirmed data base are considered to be reliable.

The decision to establish a criterion based on a non-threshold model is made after evaluating all available information on carcinogenicity and supportive information on mutagenicity. The approach and conditions for the qualitative decision of carcinogenicity are outlined in the U.S. EPA Interim Cancer Guidelines (41 FR 21402), in a report by Albert, et al. (1977), and in the Interagency Regulatory Liaison Group (IRLG) guidelines on carcinogenic risks (IRLG, 1979). It is assumed that a substance which induces a statistically significant carcinogenic response in animals has the capacity to cause cancer in humans. A chemical which has not induced a significant cancer response in humans or experimental animals is not identified as a carcinogen, even though its metabolites or close structural analogues might induce a carcinogenic response or it was shown to be mutagenic in an *in vitro* system.

It is recognized that some potential human carcinogens may not be identified by the guidelines given above. For example, compounds for which there is plausible but weak qualitative evidence of carcinogenicity in experimental animal systems (such as data from mouse skin painting or strain A mouse pulmonary adenoma) would be included in this category. The derivation of a criterion for human consumption from these studies in not valid, regardless of the qualitative outcome. In addition, there are certain compounds (e.g., nickel and beryllium) which were shown to be carcinogenic in humans after inhalation exposure by chemical form, but have induced thus far no response in animals or humans via ingesting their soluble salts. Nevertheless, a non-threshold criterion is developed for beryllium because tumors have been produced in animals at a site removed from the site of administration; in contrast, a threshold criterion is recommended for nickel because there is no evidence of tumors at sites distant resulting from administration of nickel solutions by either ingestion or injection.

For those compounds which were not reported to induce carcinogenic effects or for those compounds for which carcinogenic data are lacking or insufficient, an attempt is made to estimate a no-effect level. In many respects, the hazard evaluation from these studies is similar to that of bioassays for carcinogenicity. In order to more closely approximate conditions of human exposure, preference is given to chronic studies involving oral exposures in water or diet over a significant portion of the animal life span. Greatest confidence is placed in those studies which demonstrate dose-related adverse effects as well as no-effect levels.

There is considerable variability in the biological endpoints used to define a no-effect level. They may range from gross effects, such as mortality, to more subtle biochemical, physiological, or pathological changes. Teratogenicity, reproductive impairment, and behavioral effects are significant toxic consequences of environmental contamination. In instances where carcinogenic or other chronic effects occur at exposure levels below those causing teratogenicity, reproductive impairment, or behavioral effects, the former are used in deriving the criterion. For most of the compounds evaluated thus far, teratogenicity and reproductive impairment occur at doses near maximum tolerated levels with dose administration schedules well above estimated environmental exposure levels. Moreover, information on behavioral effects, which could be of significance, is not available for most of the compounds under study. Consequently, most NOAELs derived from chronic studies are based either on gross toxic effects or on effects directly related to functional impairment or defined pathological lesions.

For compounds on which adequate chronic toxicity studies are not available, studies on acute and subacute toxicity assume greater significance. Acute toxicity studies usually involve single exposures at lethal or near lethal doses. Subacute studies often involve exposures exceeding 10 percent of the life span of the test organism, e.g., 90 days for the rat with an average life span of 30 months. Such studies are useful in establishing the nature of the compound's toxic effects and other parameters of compound toxicity, such as target organ effects, metabolic behavior, physiological/biochemical effects, and patterns of retention and tissue distribution. The utility of acute and subacute studies in deriving environmentally meaningful NOELs is uncertain, although McNamara (1976) has developed application factors for such derivations.

In some cases where adequate data are not available from studies utilizing oral routes of administration, no-effect levels for oral exposures may be estimated from dermal or inhalation studies. Such estimates involve approximations of the total dose administered based on assumptions about breathing rates and/or magnitude of absorption.

D. Criterion Rationale

This section reviews existing standards for the chemical(s), summarizes data on current levels of human exposure, attempts to identify special groups at risk, and defines the basis for the recommended criterion.

Information on existing standards is included primarily for comparison with the proposed water quality criteria. Some of the present standards, such as those recommended by the Occupational Safety and Health Administration (OSHA) or the American Conference of Governmental Industrial Hygienists (ACGIH), are based on toxicologic data but are intended as acceptable levels for occupational rather than environmental exposure. Other levels, such as those recommended by the National Academy of Sciences in *Drinking Water and Health* (1977) or in the U.S. EPA Interim Primary Drinking Water Standards, are more closely related to proposed water quality criteria. Emphasis is placed on detailing the basis for the existing standards wherever possible.

Summaries of current levels of human exposure, presented in this section, specifically address the suitability of the data to derive water quality criteria. The identification of special groups at risk, either because of geographical or occupational differences in exposure or biological differences in susceptibility to the compound(s), focuses on the impact that these groups should have on the development of water quality criteria.

The basis for the recommended criteria section summarizes and qualifies all of the data used in developing the criteria.

IV. Guidelines for Criteria Derivation

The derivation of water quality criteria from laboratory animal toxicity data is essentially a two-step procedure: First, a total daily intake for humans must be estimated which establishes either a defined level of risk for non-threshold effects or a no-effect level for threshold effects. Secondly, assumptions must be made about the contribution of contaminated water and the consumption of fish/shellfish to the total daily intake of the chemical. These estimates are then used to establish the tolerable daily intake and consequently the water quality criterion.

A. Non-Threshold Effects

After the decision has been made that a compound has the potential for causing cancers in humans and that data exist which permit the derivation of a criterion, the water concentration which is estimated to cause a lifetime carcinogenic risk of 10^{-5} is determined. The lifetime carcinogenicity risk is the probability that a person would get cancer sometime in his or her life assuming continuous exposure to the compound. The water concentration is calculated by using the low-dose extrapolation procedure proposed by Crump (1980). This procedure is an improvement on the multistage low dose extrapolation procedure by Crump, et al. (1977).

The data used for quantitative estimates are of two types: (1) lifetime animal studies, and (2) human studies where excess cancer risk has been associated with exposure to the agent. In animal studies it is assumed, unless evidence exists to the contrary, that if a carcinogenic response occurs at the dose levels used in the study, then proportionately lower responses will also occur at all lower doses, with an incidence determined by the extrapolation model discussed below.

1. Choice of Model.

There is no really solid scientific basis for any mathematical extrapolation model which relates carcinogen exposure to cancer risks at the extremely low levels of concentration that must be dealt with in evaluating the environmental hazards. For practical reasons, such low levels of risk cannot be measured directly either using animal experiments or epidemiologic studies. We must, therefore, depend on our current understanding of the mechanisms of carcinogenesis for guidance as to which risk model to use. At the present time, the dominant view of the carcinogenic process involves the concept that most agents which cause cancer also cause irreversible damage to DNA. This position is reflected by the fact that a very large proportion of agents which cause cancer are also mutagenic. There is reason to expect that the quantal type of biological response that is characteristic of mutagenesis is associated with a linear non-threshold dose-response relationship. Indeed, there is substantial evidence from mutagenesis studies with both ionizing radiation and with a wide variety of chemicals that this type of dose-response model is the appropriate one to use. This is particularly true at the lower end of the dose-response curve; at higher doses, there can be an upward curvature, probably reflecting the effects of multistage processes on the mutagenic response. The linear non-threshold dose-response relationship is also consistent with the relatively few epidemiological studies of cancer responses to specific agents that contain enough information to make the evaluation possible (e.g., radiation-induced leukemia, breast and thyroid cancer, skin cancer induced by arsenic in drinking water, and liver cancer induced by aflatoxin in the diet). There is also some evidence from animal experiments that is consistent with the linear non-threshold hypothesis (e.g., liver tumors induced in mice by 2-acetylaminofluorene in the large scale ED_{01} study at the National Center of Toxicological Research, and the initiation stage of the two-stage carcinogenesis model in the rat liver and the mouse skin).

Because it has the best, albeit limited, scientific basis of any of the current mathematical extrapolation models, the linear non-threshold model has been adopted as the primary basis for risk extrapolation to low levels of the dose-response relationship. The risk assessments made with this model should be regarded as conservative, representing the most plausible upper limit for the risk; i.e., the true risk is not likely to be higher than the estimate, but it could be smaller.

The mathematical formulation chosen to describe the linear, non-threshold dose-response relationship at low doses is the improved multistage model developed by Crump (1980). This model employs enough arbitrary constants to be able to fit almost any monotonically increasing dose-response data and it incorporates a procedure for estimating the largest possible linear slope (in the 95 percent confidence limit sense) at low extrapolated doses that is consistent with the data at all dose levels of the experiment. For this reason, it may be called a "linearized" multistage model.

2. Procedure of Low-Dose Extrapolation Based on Animal Carcinogenicity Data.

A. Description of the Extrapolation Model

Let P(d) represent the lifetime risk (probability) of cancer at dose d. The multistage model has the form

$$P(d)=1-\exp[-(q_0+q_1d+q_2d^2+\ldots+q_kd^k)]$$

where:

$q_i>0$, and $i=0, 1, 2, \ldots, k$

Equivalently,

$$A(d)=1-\exp[-(q_1d+q_2d^2+\ldots+q_kd^k)]$$

where:

$$A(d)=\frac{P(d)-P(o)}{1-P(o)},$$

is the extra risk over background rate at dose d.

The point estimate of the coefficients q_i, $i=0, 1, 2, \ldots, k$, and consequently the extra risk function A(d) at any given dose d, is calculated by maximizing the likelihood function of the data.

The point estimate and the 95 percent upper confidence limit of the extra risk A(d) are calculated by using the computer program GLOBAL 79 developed by Crump and Watson (1979). Upper 95 percent confidence limits on the extra risk and lower 95 percent confidence limits on the dose producing a given risk are determined from a 95 percent upper confidence limit, q_1^*, on parameter q_1. Whenever $q_1\neq0$, at low doses extra risk A(d) has approximately the form $A(d)=q_1\times d$. Therefore, $q_1\times d$ is a 95 percent upper confidence limit on the extra risk and R/q_1^* is a 95 percent lower confidence limit on the dose producing an extra risk of R. Let L_0 be the maximum value of the log-likelihood function. The upper limit q_1^* is calculated buy increasing q_1 to a value q_1^* such that when the log-likelihood is again maximized subject to this fixed value q_1^* for the linear coefficient, the resulting maximum value of the log-likelihood L_1 satisfies the equation

$$2(L_0-L_1)=2.70554$$

where 2.70554 is the cumulative 90 percent point of the chi-square distribution with one degree of freedom, which corresponds to a 95 percent upper limit (one-sided). This approach of computing the upper confidence limit for the extra risk A(d) is an improvement on the Crump, et al. (1977) model. The upper confidence limit for the extra risk calculated at low doses is always linear. This is conceptually consistent with the linear nonthreshold concept discussed earlier. The slope q_1^* is taken as an upper bound of the potency of the chemical in inducing cancer at low doses.

In fitting the dose-response model, the number of terms in the polynomial g is chosen equal to (h-1), where h is the number of dose groups in the experiment, including the control group.

Whenever the multistage model does not fit the data sufficiently, data at the highest dose is deleted and the model is refitted to the rest of the data. This is continued until an acceptable fit to the data is obtained. To determine whether or not a fit is acceptable, the chi-square statistic:

$$\chi^2=\sum_{i=1}^{h}\frac{(X_i-N_iP_i)^2}{N_iP_i(1-P_i)}$$

is calculated, where N_i is the number of animals in the i^{th} dose group, X_i is the number of animals in the i^{th} dose group with a tumor response, P_i is the probability of a response in the i^{th} dose group estimated by fitting the multistage model to the data, and h is the number of remaining groups.

The fit is determined to be unacceptable whenever chi-square (X^2) is larger than the cumulative 99 percent point of the chi-square distribution with f degrees of freedom, where f equals the number of dose groups minus the number of non-zero multistage coefficients.

3. Selection and Form of Data used to Estimate Parameters in the Extrapolation Model.

For some chemicals, several studies in different animal species, strains, and sexes each conducted at several doses and different routes of exposure are available. A choice must be made as to which of the data sets from several studies are to be used in the model. It is also necessary to correct for metabolism differences between species and for differences in absorption via different routes of administration. The procedures, listed below, used in evaluating these data are consistent with the estimate of a maximum-likely-risk.

a. The tumor incidence data are separated according to organ sites or tumor types. The set data (i.e., dose and tumor incidence) used in the model is set where the incidence is statistically significantly higher than the control for at least one test dose level and/or where the tumor incidence rate shows a statistically significant trend with respect to dose level. The data set which gives the highest estimate of lifetime carcinogenic risk q_1^* is selected in most cases. However, efforts are made to exclude data sets which produce spuriously high risk estimates because of a small number of animals. That is, if two sets of data show a similar dose-response relationship and one has a very small sample size, the set of data which has the larger sample size is selected for calculating the carcinogenic potency.

b. If there are two or more data sets of comparable size which are identical with respect to species, strain, sex, and tumor sites, the geometric mean of q_1^*, estimated from each of these data sets is used for risk assessment. The geometric mean of numbers $A_1, A_2, \ldots, A_m$ is defined as $(A_1\times A_2\times\ldots\times A_m)^{1/m}$

c. If sufficient data exist for two or more significant tumor sites in the same study, the number of animals with at least one of the specific tumor sites under consideration is used as incidence data in the model.

d. Following the suggestion of Mantel and Schneiderman (1975), we assume that mg/surface area/day is an equivalent dose between species. Since to a close approximation the surface area is proportional to the ⅔rds power of the weight as would be the case for a perfect sphere, the exposure in mg/⅔rds power of the body weight/day is similarly considered to be an equivalent exposure. In an animal experiment, this equivalent dose is computed in the following manner:

Let:

L_e = duration of experiment
l_e = duration of exposure
m = average dose per day in mg during administration of the agent (i.e., during l_e)
W = average weight of the experimental animal

Then, the lifetime average exposure is

$$d=\frac{l_e\times m}{L_e\times W^{2/3}}$$

Often exposures are not given in units of mg/day, and it becomes necessary to convert the given exposures into mg/day. For example, in most feeding studies, exposure is expressed as ppm in the diet. In this case the exposure (mg/day) is derived by: $m=ppm\times F\times r$

where ppm is parts per million of the carcinogenic agent in the diet, F is the weight of the food consumed per day in kgms, and r is the absorption fraction.

In the absence of any data to the contrary, r is assumed to be one. For a uniform diet the weight of the food consumed is proportional to the calories required, which, in turn, is proportional to the surface area or the ⅔rds power of the weight, so that: $m \alpha ppm \times W^{2/3} \times r$ or

$$\frac{m}{rW^{2/3}} \alpha \; ppm$$

As a result, ppm in the diet is often assumed to be an equivalent exposure between species. However, we feel that this is not justified since the calories/kg of food is significantly different in the diet of man vs. laboratory animals, primarily due to moisture content differences. Instead, we use an empirically derived food factor, $f=F/W$, which is the fraction of a species body weight that is consumed per day as food. We use the rates given below.

Species	W	f
Man	70	0.028
Rat	0.35	0.05
Mice	0.03	0.13

Thus, when the exposure is given as a certain dietary concentration in ppm, the exposure in $mg/W^{2/3}$ is

$$\frac{m}{r \times W^{2/3}} = \frac{ppm \times F}{W^{2/3}} =$$

$$\frac{ppm \times fxW}{W^{2/3}} = ppm \times f \times W^{1/3}$$

When exposure is given in terms of $mg/kg/day = m/Wr = s$ the conversion is simply:

$$\frac{m}{rW^{2/3}} = s \times W^{1/3}$$

When exposure is via inhalation, the calculation of dose can be considered for two cases where (1) the carcinogenic agent is either a completely water-soluble gas or an aerosol and is absorbed proportionally to the amount of air breathed in, and (2) where the carcinogen is a poorly water-soluble gas which reaches an equilibrium between the air breathed and the body compartments. After equilibrium is reached, the rate of absorption of these agents is expected to be proportional to metabolic rate, which in turn is proportional to the rate of oxygen consumption, which in turn is a function of surface area.

Case 1

Agents that are in the form of particulate matter or virtually completely absorbed gases such as SO_2 can reasonably be expected to be absorbed proportional to the breathing rate. In this case the exposure in mg/day may be expressed as: $m=I\times v\times r$ where I is inhalation rate per day in m^3, v is mg/m^3 of the agent in air, and r is the absorption fraction.

The inhalation rates, I, for various species can be calculated from the observation (FASEB, 1974) that 25 gm mice breathe 34.5 liters/day and 113 gm rats breathe 105 liters/day. For mice and rats of other weights, W, (expressed in kg), the surface area proportionality can be used to determine breathing rates (in m^3/day) as follows:

For mice, $I=0.0345 (W/0.025)^{2/3} m^3/day$

For rats, $I=0.105 (W/0.113)^{2/3} m^3/day$

For humans, the values of 20 m^3/day [a] is adopted as a standard breathing rate (ICRP, 1977).

The equivalent exposure in $mg/W^{2/3}$ for these agents can be derived from the air intake data in a way analogous to the food intake data. The empirical factors for the air intake per kg per day, $i=I/W$ based upon the previously stated relationships, are as tabulated below:

Species	W	i=I/W
Man	70	0.29
Rat	0.35	0.64
Mice	0.03	1.3

Therefore, for particulates or completely absorbed gases, the equivalent exposure in $mg/W^{2/3}$ is:

$$\frac{m}{W^{2/3}} = \frac{Ivr}{W^{2/3}} = \frac{iWvr}{W^{2/3}} = iW^{1/3} vr$$

In the absence of empirical data or a sound theoretical argument to the contrary, the fraction absorbed, r, is assumed to be the same for all species.

Case 2

The dose in mg/day of partially soluble vapors is proportional to the O_2 consumption which in turn is proportional to $W^{2/3}$ and to the solubility of gas in body fluids, which can be expressed as an absorption coefficient r for the gas. Therefore, when expressing the O_2 consumption as $O_2=k W^{2/3}$, where k is a constant independent of species, it follows that $m=k W^{2/3} \times v \times r$ or

$$d = \frac{m}{W^{2/3}} = kvr$$

As with Case 1, in the absence of experimental information or a sound theoretical argument to the contrary, the absorption fraction, r, is assumed to be the same for all species. Therefore, for these substances a certain concentration in ppm or μ/m^3 in experimental animals is equivalent to the same concentration in humans. This is supported by the observation that the minimum alveolar concentration, necessary to produce a given "stage" of anesthesia, is similar in man and animals (Dripps, et al. 1977). When the animals were exposed via the oral route and human exposure is via inhalation or vice-versa, the assumption is made, unless there is pharmacokinetic evidence to the contrary, that absorption is equal by either exposure route.

e. If the duration of experiment (L_e) is less than the natural life span of the test animal (L), the slope q_1^*, or more generally the exponent g(d), is increased by multiplying a factor $(L/L_e)^3$. We assume that if the average dose, d, is continued, the age specific rate of cancer will continue to increase as a constant function of the background rate. The age specific rates for humans increase at least by the 2nd power of the age and often by a considerably higher power, as demonstrated by Doll (1971). Thus, we would expect the cumulative tumor rate to increase by at least the 3rd power of age. Using this fact, we assume that the slope q_1^*, or more generally, the exponent g(d), would also increase by at least the 3rd power of age. As a result, if the slope q_1^* [or g(d)] is calculated at age L_e, we would expect that if the experiment had been continued for the full-life span, L, at the given average exposure, the slope q_1^* [or g(d)] would have been increased by at least $(L/L_e)^3$.

This adjustment is conceptually consistent to the proportional hazard model proposed by Cox (1972) and the time-to-tumor model considered by Crump, et al. (1977) where the probability of cancer at age t and dose d is given by $P(d,t)=1-\exp[-f(t)\times g(d)]$

4. Calculation of Carcinogenic Potency Based on Human Data. If human epidemiology studies and sufficiently valid exposure information are available for the compound, they are always used in some way. If they show a carcinogenic effect, the data are analyzed to give an estimate of the linear dependence of cancer rates on lifetime average dose, which is equivalent to the factor q_1^*. If they show

[a] From "Recommendation of the International Commission on Radiological Protection," page 9, the average breathing rate is $10^7 cm^3$ per 8-hour work day and $2\times10^7 cm^3$ in 24 hours.

no carcinogenic effect when positive animal evidence is available, then it is assumed that a risk does exist but it is smaller than could have been observed in the epidemiologic study, and an upper limit of the cancer incidence is calculated assuming hypothetically that the true incidence is just below the level of detection in the cohort studied, which is determined largely by the cohort size. Whenever possible, human data are used in perference to animal bioassay data.

In human studies, the response is measured in terms of the relative risk of the exposed cohort of individuals compared to the control group. In the analysis of this data, it is assumed that the excess risk, or relative risk minus one, $R(X)-1$, is proportional to the lifetime average exposure, X, and that it is the same for all ages. It follows that the carcinogenic potency is equal to $[R(X)-1]/X$ multiplied by the lifetime risk at that site in the general population. Except for an unusually well-documented human study, the confidence limit for the excess risk is not calculated, due to the difficulty in accounting for the uncertainty inherent in the data (exposure and cancer response).

5. Calculation of Water Quality Criteria. After the value of $q_1{}^*$ in $(mg/kg/day)^{-1}$ has been determined, the lifetime risk, P, from an average daily exposure of x mg/kg/day is found from the equation $P=q_1{}^*x$. Therefore, if the lifetime risk is set at $P=10^{-5}$ for calculation purposes, the intake, I, in mg/day for a 70 kg person can be found by the equation: $I=70\times10^{-5}/q_1{}^*$ The intake of the agent from ambient water is assumed to come from two sources: (1) drinking an average of 2 liters of water per day, and (2) ingesting an average of 6.5 grams of fish per day. Because of accumulation of residues in fish, the amount of the pollutant in fish (mg/kg of edible fish) is equal to a factor R times the water concentration (mg/kg of water). Therefore, the total intake I can be written as sum of two terms: $I(mg/day)=C(mg/l)\times R(l/kg\ fish)\times0.0065\ kg\ fish/day+C(mg/l\times2l/day=C(2+0.0065R)$ where C is the water concentration in mg/l. Therefore, the water concentration in mg/l corresponding to a lifetime risk of 10^{-5} for a 70 kg person is calculated by the formula:

$$C = \frac{70 \times 10^{-5}}{q_1{}^*(2 + 0.0065\ R)}$$

B. Threshold Effects

1. Use of Animal Toxicity Data (Oral). In developing guidelines for deriving criteria based on noncarcinogenic responses, five types of response levels are considered:

NOEL—No-Observed-Effect-Level
NOAEL—No-Observed-Adverse-Effect-Level
LOEL—Lowest-Observed-Effect-Level
LOAEL—Lowest-Observed-Adverse-Effect-Level
FEL—Frank-Effect-Level

Adverse effects are defined as any effects which result in functional impairment and/or pathological lesions which may affect the performance of the whole organism, or which reduce an organism's ability to respond to an additional challenge.

One of the major problems encountered in consideration of these concepts regards the reporting of "observed effect levels" as contrasted to "observed adverse effect levels". The terms "adverse" vs. "not adverse" are at times satisfactorily defined, but due to increasingly sophisticated testing protocols, more subtle responses are being identified, resulting in a need for judgment regarding the exact definition of adversity.

The concepts listed above (NOEL, NOAEL, LOEL, LOAEL) have received much attention because they represent landmarks which help to define the threshold region in specific experiments. Thus, if a single experiment yields a NOEL, a NOAEL, a LOAEL, and a clearly defined FEL in relatively closely spaced doses, the threshold region has been relatively well defined; such data are very useful for the purpose of deriving a criterion. On the other hand, a clearly defined FEL has little utility in establishing criteria when it stands alone, because such a level gives no indication how far removed the data point is from the threshold region. Similarly, a free-standing NOEL has little utility, because there is no indication of its proximity to the LOEL, since a free-standing NOEL may be many orders of magnitude below the threshold region.

Based on the above dose-response classification system, the following guidelines for deriving criteria have been adopted:

a. A free-standing FEL is unsuitable for the derivation of criteria.

b. A free-standing NOEL is unsuitable for the derivation of criteria. If multiple NOELs are available without additional data on LOELs, NOAELs, or LOAELs, the highest NOEL should be used to derive a criterion.

c. A NOAEL, LOEL, or LOAEL can be suitable for criteria derivation. A well-defined NOAEL from a chronic (at least 90-day) study may be used directly, applying the appropriate uncertainty factor. For a LOEL, a judgment needs to be made whether it actually corresponds to a NOAEL or a LOAEL. In the case of a LOAEL, an additional uncertainty factor is applied; the magnitude of the additional uncertainty factor is judgmental and should lie in the range of 1 to 10. Caution must be exercised not to substitute "Frank-Effect-Levels" for "Lowest-Observable-Adverse-Effect-Levels".

d. If for reasonably closely spaced doses only a NOEL and a LOAEL of equal quality are available, then the appropriate uncertainty factor is applied to the NOEL.

In using this approach, the selection and justification of uncertainty factors are critical. The basic definition and guidelines for using uncertainty factors has been given by the National Academy of Sciences (1977). "Safety Factor" or "Uncertainty Factor" is defined as a number that reflects the degree or amount of uncertainty that must be considered when experimental data in animals are extrapolated to man. When the quality and quantity of experimental data are satisfactory, a low uncertainty factor is used; when data is judged to be inadequate or equivocal, a larger uncertainty factor is used. The following general guidelines have been adopted in establishing the uncertainty factors:

a. Valid experimental results from studies on prolonged ingestion by man, with no indication of carcinogenicity. Uncertainty Factor=10

b. Experimental results of studies of human ingestion not available or scanty (e.g., acute exposure only) with valid results of long-term feeding studies on experimental animals, or in the absence of human studies, valid animal studies on one or more species. No indication of carcinogenicity. Uncertainty Factor=100

c. No long-term or acute human data. Scanty results on experimental animals with no indication of carcinogenicity. Uncertainty Factor=1,000

Considerable judgment must be used in selecting the appropriate safety factors for deriving a criterion. In those cases where the data do not completely fulfill the conditions for one category and appear to be intermediate between two categories an intermediate uncertainty factor is used. Such an intermediate uncertainty factor may be developed based on a logarithmic scale (e.g., 33, being halfway between 10 and 100 on a logarithmic scale).

In determining the appropriate use of the uncertainty factors, the phrase "no indication of carcinogenicity" is interpreted as the absence of carcinogenicity data from animal experimental studies or human epidemiology. Available short-term carcinogenicity screening tests are reported in the criteria documents, but they are not used either for derivation of numerical criteria nor to rule out the uncertainty factor approach.

Because of the high degree of judgment involved in the selection of a safety factor, the criterion derivation section of each document should provide a detailed discussion and justification for both the selection of the safety factor and the data to which it is applied. This discussion should reflect a critical review of the available data base. Factors to be considered include number of animals, species, and parameters tested; quality of controls; dose levels; route; and dosing schedules. An effort should be made to differentiate between results which constitute a toxicologically sufficient data base and data which may be spurious in nature.

2. Use of Acceptable Daily Intake (ADI). For carcinogens, the assumption of low dose linearity precludes the necessity for defining total exposure in the estimation of increased incremental risk. For non-carcinogens, ADIs and criteria derived therefrom are calculated from total exposure data that include contributions from the diet and air. The equation used to derive the criterion (C) is: $C=ADI-(DT+IN)/[2\ l+(0.0065\ kg \times R)]$ where 2 l is assumed daily water consumption, 0.0065 kg is assumed daily fish consumption, R is bioconcentration factor in units of l/kg, DT is estimated non-fish dietary intake, and IN is estimated daily intake by inhalation.

If estimates of IN and DT cannot be provided from experimental data, an assumption must be made concerning total exposure. It is recognized that either the inability to estimate DT and IN due to lack of data or the wide variability in DT and IN in different states may add an additional element of uncertainty to the criterion formulation process. In terms of scientific validity, the accurate estimate of the Acceptable Daily Intake is the major factor in satisfactory derivation of water quality criteria.

3. Use of Threshold Limit Values or Animal Inhalation Studies. Threshold Limit Values (TLVs) are established by the American Conference of Governmental and Industrial Hygienists (ACGIH) and represent 8-hour time-weighted average concentrations in air that are intended to protect workers from various adverse health effects over a normal working lifetime. Similar values are set by NIOSH (criteria) and OSHA (standards) for 10- and 8-hour exposures, respectively. To the extent that these values are based on sound toxicologic assessments and have been protective in the work environment, they provide useful information for deriving or evaluating water quality criteria. However, each TLV must be carefully examined to determine if the basis of the TLV contains data which can be used directly to derive a water quality criterion using the uncertainty factor approach. In addition, the history of each TLV must be examined to assess the extent to which it has assured worker safety. In each case, the types of effects against which TLVs are designed to protect are examined in terms of their relevance to exposure from water. It must be demonstrated that the chemical is not a localized irritant and that there is no significant effect at the site of entry irrespective of the routes of exposure (i.e., oral or inhalation).

If the TLV or similar value is recommended as the basis of the criterion, consideration of the above points is explicitly stated in the criterion derivation section of the document. Particular emphasis is placed on the quality of the TLV relative to the available toxicity data that normally is given priority over TLVs or similar established values. If the TLV can be justified as the basis for the cirterion, then the problems associated with the estimation of acceptable oral doses from inhalation data must be addressed.

Estimating equivalencies of dose-response relationships from one route of exposure to another introduces an additional element of uncertainty in the derivation of criteria. Consequently, whenever possible, ambient water quality criteria should be based on data involving oral exposures. if oral data are insufficient, data from other routes of exposure may be useful in the criterion derivation process.

Inhalation data, including TLVs or similar values, are the most common alternatives to oral data. Estimates of equivalent doses can be based upon: (1) available pharmacokinetic data for oral and inhalation routes, (2) measurements of absorption efficiency from ingested or inhaled chemicals, or (3) comparative excretion data when the associated metabolic pathways are equivalent to those following oral ingestion or inhalation. Given that sufficient pharmacokinetic data are available, the use of accepted pharmacokinetic models provides the most satisfactory approach for dose conversions. However, if available pharmacokinetic data are marginal or of questionable quality, pharmacokinetic modeling is inappropriate.

The Stokinger and Woodward (1958) approach, or similar models based on assumptions of breathing rate and absorption efficiency, represents possible alternatives when data are not sufficient to justify pharmacokinetic modeling. Such alternative approaches, however, provide less satisfactory approximations because they are not based on pharmacokinetic data. Consequently, in using the Stokinger and Woodward or related models, the uncertainties inherent in each of the assumptions and the basis of each assumption must be clearly stated in the derivation of the criterion.

The use of data pertaining to other routes of exposure to derive water quality criteria may also be considered. As with inhalation data, an attempt is made to use accepted toxicologic and pharmacokinetic principles to estimate equivalent oral doses. If simplifying assumptions are used, their bases and limitations must be clearly specified.

Because of the uncertainties involved in extrapolating from one route of exposure to another and the consequent limitations that this may place on the derived criterion, the decision to disallow such extrapolation and recommend no criterion is highly judgmental and must be made on a case-by-case basis. A decision for or against criteria derivation must balance the quantity and quality of the available data against a perceived risk to the human population.

If the Stokinger and Woodward (1958) approach is used to calculate an ADI from a TLV, the general equation is: $ADI=TLV \times BR \times DE \times d \times A_A/(A_O \times SF)$ where:

ADI=Acceptable daily intake in mg
TLV=Concentration in air in mg/m^3
DE=Duration of exposure in hours per day
d=5 days/7 days
A_A=Efficiency of absorption from air
A_O=Efficiency of absorption from oral exposure
SF=Safety factor following guidelines given above
BR=Amount of air breathed per day; assume 10 m^3

For deriving an ADI from animal toxicity data, the equation is: $ADI=C_A \times D_E \times d \times A_A \times BR \times 70\ kg/(BW_A \times A_O \times SF)$ where:

ADI=Acceptable daily intake in mg
C_A=Concentration in air in mg/m^3
D_E=Duration of exposure in hours per day
d=Number of days exposed/number of days observed
A_A=Efficiency of absorption from air
BR=Volume of air breathed per day in m^3
70 kg=Assumed human body weight
BW_A=Body weight of experimental animals in kg

A_0=Efficiency of absorption from oral exposure
SF=Safety factor following guidelines given above.

More formal pharmacokinetic models must be developed on a compound-by-compound basis.

It should be noted that the safety factors used in the above formulae are intended to account for species variability. Consequently, the mg/surface area/day conversion factor is not used in the derivation of toxicity based criterion.

C. Organoleptic Criteria

Organoleptic criteria define concentrations of materials which impart undesirable taste and/or odor to water. In developing and utilizing such criteria two factors must be appreciated: the limitations of most organoleptic data and the human health significance of organoleptic properties.

The publications which report taste and odor thresholds are, with very few exceptions, cryptic in their descriptions of test methodologies, number of subjects tested, concentration: response relationships, and sensory characteristics at specific concentrations above threshold. Thus, the quality of organoleptic data is often significantly less than that of toxicologic data used in establishing other criteria. Consequently, a critical evaluation of the available organoleptic data must be made and the selection of the most appropriate data base for the criterion must be based on sound scientific judgment.

Organoleptic criteria are not based on toxicologic information and have no direct relationship to potential adverse human health effects. Although sufficiently intense organoleptic characteristics could result in depressed fluid intake which, in turn, might aggravate a variety of functional disease states (i.e., kidney and circulatory diseases), such effects are not used in the derivation process of organoleptic criteria unless available data would indicate an indirect human health effect via decreased fluid consumption, criteria derived solely from organoleptic data are based upon aesthetic qualities only.

Since organoleptic and human health effects criteria are based on different endpoints, a distinction must be made between these two sets of information. In criteria summaries involving both types of data, the following format is used:

For comparison purposes, two approaches were used to derive criterion levels for ———. Based on available toxicity data, for the protection of public health the derived level is ———. Using available organoleptic data, for controlling undesirable taste and odor quality of ambient water the estimated level is ———. It should be recognized that organoleptic data as a basis for establishing a water quality criteria have no demonstrated relationship to potential adverse human health effects.

In those instances where a level to limit toxicity cannot be derived, the following statement is to be appropriately inserted:

Sufficient data are not available for ——— to derive a level which would protect against the potential toxicity of this compound.

D. Criteria for Chemical Classes

A chemical class is broadly defined as any group of chemical compounds which are reviewed in a single risk assessment document. In criterion derivation, isomers should be regarded as a part of a chemical class rather than as a single compound. A class criterion is an estimate of risk/safety which applies to more than one member of a class. It involves the use of available data on one or more chemicals of a class to derive criteria for other compounds of the same class in the event that there are insufficient data available to derive compound-specific criteria.

A class criterion usually applies to each member of a class rather than to the sum of the compounds within the class. While the potential hazards of multiple toxicant exposure are not to be minimized, a criterion, by definition, most often applies to an individual compound. Exceptions may be made for complex mixtures which are produced, released, and toxicologically tested as mixtures (e.g., toxaphene and PCBs). For such exceptions, some attempt is made to assess the effects of environmental partitioning (i.e., different patterns of environmental transport and degradation) on the validity of the criterion. If these effects cannot be assessed, an appropriate statement of uncertainty should accompany the criterion.

Since relatively minor structural changes within a class of compounds can have pronounced effects on their biological activities, reliance on class criteria should be minimized. Whenever sufficient toxicologic data are available on a chemical within a class, a compound-specific criterion should be derived. Nonetheless, for some chemical classes, scientific judgment may suggest a sufficient degree of similarity among chemicals within a class to justify a class criterion applicable to some of all members of a class.

The development of a class criterion takes into consideration the following:

1. A detailed review of the chemical and physical properties of chemicals within the group should be made. A close relationship within the class with respect to chemical acitivity would suggest a similar potential to reach common biological sites within tissues. Likewise, similar lipid solubilities would suggest the possibility of comparable absorption and tissue distribution.
2. Qualitative and quantitative data for chemicals within the group are examined. Adequate toxicologic data on a number of compounds within a group provides a more reasonable basis for extrapolation to other chemicals of the same class than minimal data on one chemical or a few chemicals within the group.
3. Similarities in the nature of the toxicologic response to chemicals in the class provides additional support for the prediction that the response to other members of the class may be similar. In contrast, where the biological response has been shown to differ markedly on a qualitative and quantitative basis for chemicals within a class, the extrapolation of a criterion to other members of that class is not appropriate.
4. Additional support for the validity of extrapolation of a criterion to other members of a class could be provided by evidence of similar metabolic and pharmacokinetic data for some members of the class.

Based on the above considerations, it may be reasonable in some cases to divide a chemical class into various subclasses. Such divisions could be based on biological endpoints (e.g., carcinogens/non-carcinogens), potency, and/or sufficiency of data (e.g., a criterion for some members of a class but no criterion for others). While no *a priori* limits can be placed on the extent of subclassification, each subclassification must be explicitly justified by the available data.

Class criteria, if properly derived and supported, can constitute valid scientific assessments of potential risk/safety. Conversely, the development of a class criterion from an insufficient data base can lead to serious errors in underestimating or overestimating risk/safety and should be rigorously avoided. Although scientific judgment has a proper role in the development of class criteria, such criteria are useful and defensible only if they are based on adequate data and scientific reasoning. The definition of sufficient data on similarities in physical, chemical, pharmacokinetic, or toxicologic properties to justify a class criterion may vary markedly depending on the degree of structural similarity and the gravity of the perceived risk. Consequently, it is imperative that the criterion derivation section of each document in which a class criterion is recommended explicity address each of the key issues discussed above, and define, as clearly as possible, the limitations of the proposed criterion as well as the type of data needed to generate a compound-specific criterion.

A class criterion should be abandoned when there is sufficient data availabe to derive a compound-specific criterion which protects against the biological effect of primary concern; e.g., the availability of a good subchronic study would not necessarily result in the abandonment of a class criterion based on potential carcinogenicity.

The inability to derive a valid class criterion does not, and should not, preclude regulation of a compound or group of compounds based on concern for potential human health effects. The failure to recommend a criterion is simply a statement that the degree of concern cannot be quantified based on the available data and risk assessment methodology.

E. Essential Elements

Some chemicals, particularly certain metals, are essential to biological organisms at low levels but may be toxic and/ or carcinogenic at high levels. Because of potential toxic effects, it is legitimate to establish criteria for such essential elements. However, criteria must consider essentiality and cannot be established at levels which would result in deficiency of the element in the human population.

Elements are accepted as essential if listed by NAS Food and Nutrition Board or a comparably qualified panel. Elements not yet determined to be essential but for which supportive data on essentiality exists need to be further reviewed by such a panel.

To modify the toxicity and carcinogenicity based criteria, essentiality must be quantified either as a "recommended daily allowance" (RDA) or "minimum daily requirement" (MDR). These levels are then compared to estimated daily doses associated with the adverse effect of primary concern. The difference between the RDA or MDR and the daily doses causing a specified risk level for carcinogens or ADIs for non-carcinogens defines the spread of daily doses from which the criterion may be derived. Because errors are inherent in defining both essential and maximum tolerable levels, the criterion is derived from dose levels near the center of such a dose range. The decision to use either the MDR or RDA is guided by the spread of the doses and the quality of the essentiality and toxicity estimates.

The modification of criteria by consideration of essentiality must take into account all routes of exposure. If water is a significant source of the MDR or RDA, the criterion must allow for attainment of essential intake. Conversely, even when essentiality may be attained from nonwater sources, standard criteria derivation methods may be adjusted if the derived criterion represents a small fraction of the ADI or MDR. On a case-by-case basis, the modification in the use of the guidelines may include the use of different safety factors for non-carcinogens or other modifications which can be explicitly justified.

F. Use of Existing Standards

For some chemicals for which criteria are to be established, drinking water standards already exist. These standards represent not only a critical assessment of literature, but also a body of human experience since their promulgation. Therefore, it is valid to accept the existing standard unless there is compelling evidence to the contrary. This decision should be made after considering the existing standards vs. new scientific evidence which has accumulated since the standards have been established. There are several instances where the peer review process recommended usage of the present drinking water standards.

Bibliography

Albert, R.E., et al. 1977. Rationale developed by the Environmental Protection Agency for the assessment of carcinogenic risks. Jour. Natl. Can. Inst. 58: 1537.

Cox, C.R. 1972. Regression model and life tables. Jour. Roy. Stat. Soc. B. 34: 187.

Crump, K.S. 1979. Dose-response problems in carcinogenesis. Biometrics. 35: 157.

Crump, K.S. 1980. An improved procedure for low-dose carcinogenic risk assessment from animal data. Jour. Environ. Path. Toxicol. (In press).

Crump, K.S. and W.W. Watson. 1979. GLOBAL 79. A FORTRAN program to extrapolate dichotomous animal carcinogenicity data to low dose. Natl. Inst. Environ. Health Sci. Contract No. 1-ES-2123.

Crump, K.S., et al. 1977. Confidence intervals and test of hypotheses concerning dose-response relations inferred from animal carcinogenicity data. Biometrics. 33: 437.

Doll, R. 1971. Weibull distribution of cancer: Implications for models of carcinogenesis. Jour. Roy. Stat. Soc. A. 13: 133.

Dripps, R.D., et al. 1977. Introduction to Anesthesia, the Principles of Safe Practice. 5th ed. W.B. Saunders Co., Philadelphia, Pennsylvania. p. 121.

Druckrey, H. 1967. Quantitative Aspects of Chemical Carcinogenesis. *In:* R. Turhaut (ed.), Potential Carcinogenic Hazards from Drugs, Evaluation of Risks. UICC Monograph Series, Vol. 7. Springer Verlag, New York.

Federation of American Societies for Experimental Biology. 1974. Biological Data Books, 2nd ed. Vol. III. P.L. Altman and D.S. Dittment (eds.) Federation of American Societies for Experimental Biology. Bethesda, Maryland. Library of Congress No. 72-87738.

Guess, H., et al. 1977. Uncertanity estimates for low dose rate extrapolations of animal carcinogenicity data. Can. Res. 37: 3475.

Interagency Regulatory Liaison Group. 1979. Scientific bases for identifying potential carcinogens and estimating their risks. February 6.

International Commission on Radiological Protection. 1977. Recommendation of the ICRP, Publ. No. 26, adopted January 17, 1977. Pergamon Press, Oxford, England.

Mantel, N. and M.A. Schneiderman. 1975. Estimating "safe lavels," a hazardous undertaking. Can. Res. 35: 1379.

McNamara, B.P. 1976. Concepts in Health Evaluation of Commercial and Industrial Chemicals. *In:* M.A. Mehlman, et al. (eds.), Advances in Modern Toxicology, Vol. 1, Part 1. Wiley and Sons, New York. p. 61

National Academy of Sciences. 1977. Drinking Water and Health. National Academy of Sciences, Washington, D.C.

National Cancer Institute. 1976. Guidelines for carcinogen bioassay in small rodents. NCI Carcinogenesis Technical Report Series, No. 1, February, 1976.

Peto, R. 1973. Weibull distribution for continuous-carcinogenesis experiments. Biometrics. 29: 457.

Stephan, C.E. 1980. Memorandum to J. Stara. U.S. EPA. July 3.

Stokinger, M.E. and R.L. Woodward. 1958. Toxicologic methods for establishing drinking water standards. Jour. Am. Water Works Assoc. 50: 517.

U.S. EPA. 1976. National Interim Primary Drinking Water Regulations. Office of Water Supply. EPA 570/9-76-003. U.S. Environ. Prot. Agency.

U.S. EPA. 1980. Seafood consumption data analysis. Stanford Research Institute International, Menlo Park, California. Final report, Task 11, Contract No. 68-01-3887.

Veith, G.D. 1980. Memorandum to C.E. Stephan. U.S. EPA April 14.

Veith, G.D., et al. 1979. Measuring and estimating the bioconcentration factors of chemicals in fish. Jour. Fish Res. Board Can. 36: 1040.

Health Effects Methodology—Authors, Contributors and Reviewers

Jerry F. Stara, ECAO-Cin, U.S. Environmental Protection Agency
Elizabeth L. Anderson, OHEA, U.S. Environmental Protection Agency
Roy E. Albert, N.Y. University Medical School, New York, NY
Patrick R. Durkin, Syracuse Research Corporation, Syracuse, NY
Steven D. Lutkenhoff, ECAO, U.S. Environmental Protection Agency
Robert E. McGaughy, CAG, U.S. Environmental Protection Agency
Charles Ris, OHEA, U.S. Environmental Protection Agency
Lawrence Anderson, CAG, U.S. Environmental Protection Agency
Dolph Arnicar, CAG, U.S. Environmental Protection Agency
Steven Bayard, CAG, U.S. Environmental Protection Agency
Edward Calabrese, University of Massachusetts, Amherst, MA
Thomas Clarkson, University of Rochester, Rochester, NY

Joseph Cotruvo, ODW, U.S. Environmental Protection Agency
Christopher DeRosa, ECAO-Cin, U.S. Environmental Protection Agency
Kurt Enslein, Genesee Computer Center, Inc., Syracuse, NY
Jeffrey Gaba, OGC, U.S. Environmental Protection Agency
Frank Gostomski, OWRS, U.S. Environmental Protection Agency
Daniel Greathouse, HERL-Cin, U.S. Environmental Protection Agency
Joseph Arcos, Tulane University, New Orleans, LA
Dev Barnes, OWRS, U.S. Environmental Protection Agency
David L. Bayliss, CAG, U.S. Environmental Protection Agency
Chao W. Chen, CAG, U.S. Environmental Protection Agency
Herbert Cornish, University of Michigan, Ann Arbor, MI
Kenneth Crump, Science Research Systems, Inc., Ruston, LA
Michael Dourson, ECAO-Cin, U.S. Environmental Protection Agency
John R. Fowle III, CAG, U.S. Environmental Protection Agency
R. John Garner, HERL-Cin, U.S. Environmental Protection Agency
Terrence Grady, ECAO-Cin, U.S. Environmental Protection Agency
Bernard Haberman, CAG, U.S. Environmental Protection Agency
Paul Hammond, University of Cincinnati, Cincinnati, OH
Charalingayya Hiremath, CAG, U.S. Environmental Protection Agency
Dinko Kello, Institute for Medical Research and Occupational Health, Zagreb, Yugo.
Chang S. Lao, CAG, U.S. Environmental Protection Agency
Marvin Legator, University of Texas Medical Branch, Galveston, TX
Jack McGinnity, OTS, U.S. Environmental Protection Agency
Debdas Mukerjee, ECAO-Cin, U.S. Environmental Protection Agency
Ellen O'Flaherty, University of Cincinnati, Cincinnati, OH
Fred Passman, Energy Resources Co., Inc. Cambridge, MA
John F. Risher, ECAO-Cin, U.S. Environmental Protection Agency
Jeffery Rosenblatt, OTS, U.S. Environmental Protection Agency
Samuel Shibko, US FDA, Washington, DC
Harry T. Skalsky, Reynolds Metal Company
Todd W. Thorslund, CAG, U.S. Environmental Protection Agency
Benjamin VanDuuren, N.Y. University Medical School, New York, NY
James R. Withey, National Health and Welfare Canada, Ottawa, Canada
Rolf Hartung, University of Michigan, Ann Arbor, MI
Rudolf Jaeger, Harvard School of Public Health, Cambridge, MA
Arnold M. Kuzmack, CAG, U.S. Environmental Protection Agency
Si Duk Lee, ECAO-Cin, U.S. Environmental Protection Agency
Leland McCabe, HERL-Cin, U.S. Environmental Protection Agency
Myron Mehlman, Mobil Oil Corporation, New York, NY
Jean Munson, University of Cincinnati, Cincinnati, OH
Nancy Othmer, OGC, U.S. Environmental Protection Agency
David J. Reisman, ECAO-Cin, U.S. Environmental Protection Agency
Paula K. Roberson, U.S. Environmental Protection Agency
H. Daniel Roth, Roth Associates
Dharm V. Singh, CAG, U.S. Environmental Protection Agency
Robert G. Tardiff, National Academy of Sciences, Washington, DC
Anne W. Trontell, Energy Resources Co., Inc., Cambridge, MA
John Van Ryzin, Columbia University, New York, NY
Ronald E. Wyzga, Electric Power Research Institute

Appendix D—Response to Comments on Guidelines for Deriving Water Quality Criteria for the Protection of Aquatic Life and Its Uses

Introduction

Two versions of the Guidelines were published in the Federal Register for comment. The first version (43 FR 21506, May 18, 1978 and 43 FR 29028, July 5, 1979) was simply published for comment. The second (44 FR 15926, March 15, 1979) was published as part of the request for comments on the water quality criteria for 27 of the 65 pollutants. The second version was meant to be clearer and more detailed than the first, but very similar technically. Since the two versions were so similar, comments on both will be dealt with simultaneously.

Many comments were received that no draft water quality criteria for any of the 65 pollutants should have been issued for public comment until the comments on the first version of the Guidelines had been dealt with adequately and the Guidelines changed appropriately. The comments on the first version were read and the Guidelines were revised in an attempt to make the second version clearer and more detailed than the first. However, an extensive revision of the technical content of the Guidelines was not attempted between the first and second versions because the Agency was preparing water quality criteria based on the Guidelines. The Agency could have avoided this criticism simply by not publishing any version of the Guidelines for comment until March 15, 1979, but this would have greatly reduced the length of time available for people to consider the Guidelines and comment on them. As it was, some people commented that the comment period announced on March 15, 1979, was too short.

1. Comment—The procedures used to derive criteria in the "Red Book" were upheld in court and probably should still be used.

Response—The procedures used in the Guidelines are similar to some of the procedures used to develop criteria in the "Green Book", "Blue Book", and "Red Book". The Guidelines are designed to be more objective and systematic, to deal more adequately with residues, and to incorporate the concept of a minimum data base.

2. Comment—Criteria should be compilations of critically reviewed data with no synthesis or interpretation.

Response—Neither P.L. 92-500 nor the Consent Decree specify the form which a criterion must take. The Consent Decree (para. 11, p. 14) specifies that such criteria "shall state, *inter alia*, recommended maximum permissible concentrations". Adequate precedents have been set in the "Green Book", "Blue Book", and "Red Book" for the form of criteria used in the Guidelines.

3. Comment—The Guidelines and criteria should be developed by a consensus of aquatic toxicologists rather than by EPA personnel only.

Response—EPA certainly wants the Guidelines and the criteria to be as good as possible and as acceptable to as many interested people as possible. To this end, EPA has widely distributed draft versions of the Guidelines and the criteria documents, discussed them with many people, considered the comments received, and made many significant technical changes and editorial revisions. It is questionable whether or not a true consensus could have been reached by any means within the time available. In addition, EPA has a legislative responsibility which it should not delegate to someone else.

4. Comment—The Guidelines should be updated regularly.

Response—The Guidelines are not being promulgated as a regulation or directive. The purpose of presenting these Guidelines is to show how the water quality criteria for aquatic life were derived for the 65 pollutants. If EPA uses these Guidelines again, they will be revised to take into account new data, concepts, and ideas.

5. Comment—The objectives, purpose, and limitations of the Guidelines should be stated.

Response—The introductory portion of the Guidelines has been expanded to address these subjects more fully.

6. Comment—The Guidelines are too ambiguous.

Response—The Guidelines have been revised and rewritten, partly to improve clarity and provide additional details. It is not possible to provide explicit details on all items; in some areas only general guidance can be provided at this time.