ATTACHMENT II

CMA EXHIBIT A

EVALUATION OF BIOLOGICAL DATA USED IN DERIVING EPA WATER QUALITY CRITERIA

(as stated in the May 18, 1978 Federal Register Vol. 43, No. 97, pp. 21506-21518)

by

John Cairns, Jr.

Arthur L. Buikema, Jr.

James G. Geiger

of

Virginia Polytechnic Institute and State University

Prepared for:

American Petroleum Institute

August 2, 1978

The following comments were based on review of approximately eighty-five percent of the available literature. Some personal communications were missing or incomplete, and certain references were not available for use. Since some of the data cited in the document were reordered, recalculated, or otherwise used, it was sometimes difficult to trace their origins. Consequently some of the comments herein may not be applicable. In this report, references (numbers in parentheses) are citations specific to each table.

SUMMARY OF FINDINGS

The following summary is based upon a review of the proposed criteria development methodology document and approximately eighty-give percent of the references cited therein.

MATC

The proposed MATC calculations are too restrictive. The criteria preclude using actual MATC data even if the actual MATC data prove that a toxicant is less harmful than predicted by calculation.

Inadequate Chronic Data Base

The guidelines of the criteria methods indicate that no criterion should be recommended without chronic values and acute continuous flow tests where actual toxicant concentrations are measured in the test chambers. Chronic tests require a long time, and few facilities for chronic testing exist. Since few chronic studies have been repeated, we question the reproducibility of chronic test data. Chronic studies are subject to many variables not usually experienced in acute tests--for example, seasonal and shorter cycle temperature variations, changes in test water quality over the period of exposure, and feeding and nutritional problems. Chronic test values have greater variability than do acute tests, due to problems associated with larval viability, spawning times, egg production rates, hatchability rates, and the like.

We feel that chronic data can be used to establish water quality criteria, provided additional efforts to confirm reported chronic results are carried out.

Different Water Characteristics

The criteria method development document does not meet its stated objective, namely, to "provide a basis for deriving

criteria specific to different water characteristics for pollutants whose toxicity varies significantly under different conditions." This is based on the obvious bias for soft-water data (and other factors), as indicated by the selection of data from cited references. If the criteria development methods were set up to meet this objective, the design methods and rationale were not presented. We therefore cannot comment on the methods and applications relative to the stated objective.

Water Quality Interactions

Variations in such physical parameters as temperature and in such chemical parameters as pH and hardness affect the sensitivity of organisms to many pollutants. Yet the data used to develop static vs. flow-through, nominal vs. measured, and time of exposure correction factors were selected under a wide range of temperature and water quality regimes. These data should not be grouped to develop a single correction factor under the above physical and chemical variables.

The correction factor for sensitivity does not allow for water quality interactions that affect organism sensitivity to toxicants. "More information is needed about the effects of water quality in chronic toxicity for the sake of extrapolating from one exposure condition to another." (Andrews, et al., unpublished.)

Identification of Data

As noted above, much of the data in the tables was difficult to trace to original references, because the data had been recalculated and reordered (especially references in Table 4, #42 a-o). Additionally, a number of transcription errors in our partial review of cited references were identified.

Inappropriate Units

Data are unitless throughout the tables. We found that in most of the references, data were reported as mg/l, but the units for some data included in the tables were actually µg/l. These unitless values were treated by EPA as numerical equivalents in data manipulations, resulting in inaccurate correction factors. These factors should be recalculated.

Data Selection

Some data were selected from referenced citations. Other relevant data in the same citations were not used--a practice which may well bias the correction factors. The rationale for this partial selection should have been presented.

Most of the data used to develop the correction factors were based on pesticide toxicity. This probably skews the cor-

rection factors toward pesticides.

Data presented in the tables were selected from predominantly soft-water conditions. Much of the hardwater data in the same references were not used. Again, this selection biases the correction factors, since some toxicants, especially heavy metals, are more toxic to organisms in soft water than in hard water.

LC_{50} Determination

LC_{50} values in the references cited were calculated by different methods, such as graphic interpolation (which has no confidence limits), computerized probit analysis (several methods), and moving averages. The data should all be calculated by the same method so as to reduce error in subsequent correction factor calculations.

LC_{50} and EC_{50} Values

LC_{50} and EC_{50} values were treated as numerical equivalents in all calculations. They are not measures of the same thing. EC_{50} values are measures of immobilization of an organism, whereas LC_{50} values are measures of death. Not only is immobilization more difficult to quantify accurately, but also it may be temporarily followed by subsequent animal recovery. EC_{50} values being more restrictive than LC_{50} values, use of EC_{50} data resulted in correction factors that are too stringent.

Species Sensitivity

As expected, different species have different sensitivities to toxicants. This is due to different habitat and environmental requirements, physiological requirements, age and size differences, reproductive cycles, and other factors. We suspect that it is invalid to lump all fish data and all invertebrates data together to determine correction factors.

Acclimation of Test Species

Improper acclimation of test species may modify the sensitivity of organisms to toxicants. Some references that we examined did not specify the source of the test organisms, and others specified field populations that may have been stressed by pollutants in the natural habitat. Still other references did not state their methods for acclimating organisms to test conditions prior to toxicant exposure. We therefore do not feel that these data should be used for correction factor calculations.

Test Methodology

Data collected by static, static-with-renewal, and continuous-

flow techniques probably will not give the same LC_{50} values. In table 4, data from all three exposure methods were used to generate correction factors. This is inappropriate.

RECOMMENDATIONS

1. EPA should conduct an experiment specifically designed to test the validity of their correction factors. The experiment should be designed with the aid of a statistician and should include simultaneous tests on the following variables: different species; distinctly different classes of compounds and various states of these compounds that would be found in the environment; progressive lengths of exposure, for example, readings at 24, 48, 72, and 96 hours; different physical/chemical water quality regimes; and test conditions, such as static flow-through. Analysis of variance and covariance (see Wickramarartne) should be used to establish and correct for these interrelationships prior to any calculations of correction factors.

2. Actual MATC values should be used when they exist. If actual MATC data are excluded, the reason for exclusion should be stated.

3. A more balanced data pool--that is, a pool not so heavily skewed to pesticides--should be used to calculate correction factors. EPA should not use data where the history of origin and acclimation of test organisms are not documented or are questioned. These data should be eliminated prior to recalculation.

4. EPA should add the data in the cited references which were not used in calculations and as well should seek additional data to balance soft-water and hard-water conditions in order to approximate the range of water qualities in the United States. Appropriate statistical methods should then be used to quantify these interactions. The same comment applies to pH, temperature, and type of toxicant (based on its mode of action, solubility, ionic or chelation state, and the like).

5. All data in the tables should be compared to the original literature sources, and the values should be corrected prior to recalculation of geometric means and variance (or, preferably, regression analysis, analysis of variance, and analysis of covariance. (see Wickramarartne). The data should all be converted to consistant units before any calculations are made.

6. Only LC_{50} data with stated confidence limits should be used in any calculations. All values derived by graphic interpolation should be eliminated prior to recalculation.

7. Quotients of EC/LC_{50} or EC_{50} should be recalculated. Quotients should be based on LC_{10} or EC_{10} divided by the LC_{50} or EC_{50}, respectively. The LC_{10} and EC_{10} data base should be used to compensate for random death of test organisms that occurs under experimental conditions. Where LC_{10} and EC_{10} data are not available, then we suggest this value be calculated and used to correct for random death. This approach would broaden the data base and affect the geometric mean.

8. The criteria development document should (a) clearly reference the source of data points; (b) indicate how and why data were recalculated prior to inclusion in the tables; and (c) state which criteria were used in partial selection of data from the original references.

9. The document should include a statement about how the "basis for developing criteria specific to different water characteristics" was determined.

SPECIFIC COMMENTS

Page 21507

It is true, as stated, that LC_{50} values are not always comparable; but there are many factors in addition to those named, length of exposure concentration, and type of test system. The discussion in the document ignores the interaction of the more obvious physical/chemical parameters such as temperature, pH, water hardness (and other water quality factors), and the effect of previous acclimation and history on test organisms.

Temperature influences organism sensitivity, and most organisms are more sensitive to toxicants at higher temperatures. Toxicities of trace metals, ammonia, certain phenolics, and some pesticides are affected by variations in pH and/or water hardness. These interactions are ignored throughout all correction factor calculations. Examination of cited data showed that temperatures ranged from 8° to 28°C and hardness values from <40 to >300ppm.

Table 1

No comments.

Table 2

Quotients presented may reflect data which would invalidate their use. Specifically:

(1) Temperature

(a) Values ranged from 12° to 28°C, as determined by examination of selected references in the table; and

(b) Several papers did not state test temperatures. (References 2, 5, 6, 8, 12, 13.)

(2) Hardness

The majority of data were selected for "soft water" (generally a worst case), that is, hardness <60ppm. A significant amount of data for well-defined "hard water" in these same references were ignored. This practice skews the data to a "worst case" situation and biases the correction factor. More importantly, it is inconsistent with the stated objective, of providing a "basis for deriving criteria specific to different water conditions." We should call attention to the use of one reference (19) which has several toxicity values reflecting hard-water conditions (300mg/l).

(3) Water Quality Data

In one paper examined (reference 14), no water chemistry data were presented. (We did not examine all references.) Without assurance that water quality data meet strict EPA bioassay guidelines, the toxicity data should not be used in calculations of correction factors.

(4) Calculations

LC_{50} and EC_{50} values were computed in several ways--Litchfield-Wilcoxon, computer probit analyses, graphical interpolation, and so on. It is not appropriate to combine these values in correction factor calculations. They should all be converted to a standard calculation. Graphic interpolation values, especially, either should be converted to a stronger calculation that has stated confidence intervals, or should not be used.

(5) Incomplete Data Usage

Data existed for some additional compounds in the references EPA cited, but these values were not used. For example,

Reference	Compound
6	Toxaphene
4	101 compounds (only one cited)
10	Mercury

For several papers, 24-, 48-, or 72-hour LC_{50} values were not reported, yet quotients were reported (Table references 3, 11, 33, 35, 38).

References 7 and 15 contained information for certain fish species which were not used. In reference 11, no LC_{50} values for mercuric chloride were identified. We presume the LC_{50} value for methylmercuric chloride was used.

Table 3 Data in Table 3 (of those references we verified) were based on toxicity studies conducted on five species of fish. Experimental conditions varied as follows:

Temperature -- 17° to 25°C,
Hardness -- assumed 40 to 210mg/l, and
Calculations -- various methods, including graphic interpolation.

Please see information in Table 2 for discussion of these factors.

Three notable errors were identified.

(1) Reference 4

In the nickel study, one 96-hour LC_{50} value, 27 mg/l, was a nominal--not measured--concentration. The ratios developed in the water quality document of continuous flow to static flow data were derived from 28/32 = 0.88 and 25/27 = 0.93. It is not clear whether both the static and flow-through tests were run simultaneously. We could conceivably calculate 25/32 = 0.78 and 28/27 = 1.04 based on the quoted reference and increase the range of the ratios.

(2) Reference 5

The static 96-hour LC_{50} value of 9 mg/l came from a different study, and it was divided into an arithmetic mean value of 10.45 mg/l. If the data in Table 7 of reference 5 are interpreted correctly, the bluegill data were not included in the water quality criteria document. In this instance, the ratios obtainable for bluegills could range from 0.77 to 1.31, in contrast to the 1.16 reported for the fathead minnow.

(3) Reference 6

The 96-hour LC_{50} value for endrin (0.39) was given in the text, but it is not clear whether the value came from research presented in the paper (6), other unreported research, or another paper. Reference 6 suggests that fathead minnow sensitivity to endrin varied by a threefold factor when comparing various studies. Furthermore, it is not clear whether theoretical concentrations of endrin were used to estimate LC_{50} values.

In conclusion, most of the data are for pesticides and may introduce a bias in the correction factor for other compounds.

able 4 Most of the data, especially the last sections, were difficult to comprehend due to the method of citation. This was especially true for reference 42, and where there were many citations and the numbers were rearranged.

The data might be comparable for any given toxicant or citation. It would be hard to acknowledge, however, that all data in Table 4 taken as a whole are comparable. In light of the data we examined, we found:

(1) Physical/Chemical Parameters (see Table 2 for a discussion of significance)

Temperature -- 10° to 26°C (4, 13)
Hardness -- soft to 375 mg/l (38)
pH -- varies with hardness

(2) Biological Parameters

The data base consisted of at least 20 species of fish (12) and different sizes or ages of fish (4).

(3) Testing Procedures

(a) Calculations -- different methods
(b) Conditions -- static, flow-through, and static-renewal
(c) Concentrations -- some values were nominal not measured (39)
(d) Solvent controls -- not reported (39)

(4) Units of Reporting

Not all values were corrected to mg/l, and some data were reported as µg/l (21, 29). These values were not corrected to mg/l before the geometric mean was calculated.

(5) Incipient Values

Incipient LC_{50} values were used for comparative purposes when 96-hour values were available (23). LC_{50} incipient values are not the same as standard acute tests because the time interval on an incipient value is not fixed. The 7-day incipient lethal value is 0.009 mg/l endosulfan, and it is reported as 0.09.

:able 5 We had difficulty determining how some MATC values were calculated from the cited literature. Not all of the MATC values were in the same units (mg/l and µg/l). For examples, see reference 15 and 18.

The authors of reference 15 concluded that the safe concentration of HCN for bluegills was 5.2 mg/l. They arrived at this value because 5.2 mg/l inhibited spawning behavior, and if a "complete life history is to be accomplished in cyanide-polluted waters," the level must be lower. The water quality criteria document uses 17.4 as the MATC from reference 15.

Table 6 — The general comments in Table 2 apply to Table 6. Errors in data transcription are noted below.

Reference	Criteria Citation	Correct Value
1	0.48	2.06
1	0.67	0.94
2	0.31	0.003
2	0.97	0.097
2	0.40	0.04
2	0.76	0.076
4	0.43	0.57
5	0.01	0.007
5	0.01	0.004
5	0.65	0.065
5	0.22	0.25
6	0.43	0.13
6	0.85	0.085
6	0.03	0.05
7	0.06	0.6

The correction factors should be recalculated. We were unable to check the values in reference 3.

Table 7 — We have no problems with the data presented. However, since all data used in this table are based on pesticides, the correction factor will be biased against all other compounds.

Table 8 — Those data we could trace were accurate. The upper part of the table was inconsistent in units used (mg/l and µg/l). We assume that these were not corrected prior to correction factor calculation.

Errors in data transcription include reporting a 24-hour TL_m (1). The values for zinc, nickel, and cadmium were for snail eggs, adult snails, or both (1). For thiodan (13), the value should be 75 and not 175. Data for Daphnia magna were reported in reference 13, but they were not included in the document. The temperature range used in reference 13 was from 8° to 19°C. This table includes fish data even though it is labelled for invertebrates (reference 2). For copper, EPA cites data from references 8 and 9; we could not identify the copper value for reference 8.

ble 9 The MATC value for toxaphene could not be traced. If we assume that it was from reference 42F, then no value was calculated or implied in the reference. Again MATC values were presented in mg/l and µg/l, and we assume that calculations were made without correction to comparable units.

ble 10 Only references 3 and 9 were available for study. Again, many of the comments discussed above apply, for example, lumping of data sets for species, habitat, different LC_{50} methodologies, and the like. For reference 3, 4.8 was the 96-hour LC_{50} for LAS obtained from a static test where the concentrations were calculated, not measured. For the continuous-flow test, the 96-hour LC_{50} value was 4.35, and this value was measured. The document implied that the data from reference 9 are on an effluent, but they are for pure industrial chemicals. There were errors in reporting the data. Line 6 (p. 21518) should be 0.25/0.55 = 0.45; line 15 (p. 21518) should be 5000/1000 = 0.5; and line 16 (p. 21518) should be 3200/6000 = 0.53. The table includes quotients of the highest concentration effecting or killing 0 - 10% of the organisms divided by the 96-hour LC_{50} or EC_{50} value. Where 0 and 10% mortality data were available, the data were selected to include only the highest concentration killing 0% of the population; data where there was a 10% kill (or random death) were ignored. For example, see page 21517, column 3, lines 73 (25/40=0.63), 76 (7/11 = 0.64) 77 (450/550 = 0.82), and 83 (180/290 = 0.62). Pertinent data where there was only a 10%, and not 0%, mortality reported (see tables of reference 9) were ignored. The 0.44 mean value was calculated for data where there was no observed effect or an observed 10% effect. Whey could not a calculated LC_{10} value be used to obtain the same information?