Technical Report

No 56

Aquatic Toxicity Data Evaluation

December 1993

ISSN-0773-8072-56

Technical Report No. 56

Aquatic Toxicity Data Evaluation

December 1993

ISSN-0773-8072-56

Brussels, December 1993

ECETOC Technical Report No. 56

AQUATIC TOXICITY DATA EVALUATION

CONTENTS

SUMMARY

A database, termed the ECETOC Aquatic Toxicity (EAT) database, consisting of original published information on the toxicity of substances to aquatic species in fresh and saline waters has been compiled. The principal quality criteria for acceptance of data were that test methods should be well described and the toxicant concentrations must be measured. On this basis 42% of the 530 papers examined were found to be suitable for inclusion in the EAT database.

The EAT database input software is easy to use and has been prepared for Personal Computers operating in a DOS environment. For each entry there are 21 fields of information on the substances, test species, test details, results and source references. All the references are held at ECETOC.

The EAT database includes information on 368 substances for 122 aquatic test species. Publications from 1970 to 1991 have been assessed, giving 2200 entries; organohalogens and heavy metals make up most of the data entries which reflects historical concern over these chemicals. The toxicity of all substances was evenly distributed in a log normal scale. Some groups of substances were found to be more toxic than others. The higher toxicity is not necessarily linked with substances of historical concern. Sensitivity of the test organisms to chemicals was evenly distributed. On the scale of sensitivity bacteria seem to be least sensitive and invertebrates most sensitive as a general rule.

Analysis of the data can be performed using a number of specially written routines. These include the ability to select data using various options, to prepare simple counts, frequency distributions, ratios (e.g. between acute and chronic results) and regression analyses. The regression analyses take account of the fact that in comparing toxicity test end-points there are no dependent or independent axes in the strictest sense.

In order to provide a scientific basis for application factors used in risk assessment, the ratio of acute EC_{50}:chronic NOEC was assessed for 12 different selections of data. The median ratios varied from 3.6 to 28.0. When 19 substances typical of those which could be notified under the provisions of the Seventh Amendment Directive (92/32/EEC) were considered the range of ratios was 1.25 to 28.3. The maximum acute EC_{50}:chronic NOEC value of 28.3 (for 100% of substances) is an indication that the factor of 40 (for 90% of substances) given in ECETOC (1993) may be rather conservative. The latter was derived from the same database using a different statistical approach, i.e. not allowing for the separate assessment of individual species.

This report includes a summary of approaches to hazard assessment, especially the use of different application factors and describes the result of these using three substances.

SECTION 1. INTRODUCTION

In response to the continuing demands of public, government and industry to safeguard the aquatic environment while maintaining effective and beneficial use of industrial chemicals, there is an increasing need to carry out a risk assessment of substances. A multiplicity of approaches has evolved in an attempt to meet this need.

Aquatic risk assessment requires different approaches from methods used to protect man. The protection of the aquatic environment is essentially a matter of protecting populations and their habitats. In order to do this relatively small numbers of individuals belonging to species representing the taxonomic group of concern are studied in the laboratory or in model ecosystems. In man the emphasis must be on the protection of the individual using other species to acquire the experimental data. In the safeguarding of the aquatic environment the species of concern may be examined directly, but results from a relatively small number of individuals are used to protect populations and ecosystems. Furthermore, it is necessary to extrapolate from the largely artificial nature of laboratory tests to conditions in the environment at large and, in particular, to consider the bioavailability of the substance in the environment. In carrying out aquatic risk assessment and calculations of acceptable concentrations, it has to be taken into consideration that some important industrial substances also occur naturally in surface waters, some of them even being essential for aquatic life.

In response to the need to establish methods for testing substances to define their potential for harm or hazard to the aquatic environment, a compromise has evolved between the very large testing programmes needed to acquire maximum confidence and the need to develop reasonably comprehensive but cost-effective programmes. It is recognised that the toxicity of a given compound to aquatic organisms may be influenced by a range of biological, chemical and physical factors. These include, for example, the age of the test organism, the dilution water characteristics and the test temperature. Consideration has therefore to be given to these experimental factors when evaluating aquatic toxicity data. Such considerations are included in regulatory protocols for aquatic toxicity testing which by standardisation seek to limit these confounding factors. The OECD has led these efforts internationally, producing, since 1979, guidelines for the testing of substances (OECD, 1981). The OECD guidelines have been adopted to a great extent by the European Community.

Guidelines and protocols certainly contribute to improved reproducibility of test results which has been of great importance when the objective was only to classify substances on their inherent

properties. Reproducibility was achieved at the cost of reduction in ecological relevance. With the move from hazard identification to risk assessment there is a need to bridge the gap between laboratory and field and this requires an examination of non-standard species and test conditions. The EAT database established herein draws on the literature from all types of aquatic toxicity assessment and allows an analysis of broader issues based on limited experience.

With the above in mind the Aquatic Toxicity Data Evaluation Task Force was formed initially to consider the potential toxicity of substances in the aquatic environment, the evaluation of species used for testing and the relationships between acute and chronic toxicity for aquatic organisms.

The Terms of Reference were:

- collect and review data on substances which have been tested for their effects on aquatic organisms in acute, sub-chronic and chronic tests;
- comment upon the interpretation of data on the acute toxicity and their relevance to the overall hazard posed by substances;
- examine the relationships between the results obtained from different species for acute, sub-chronic and chronic tests;
- recommend a strategy for assessing risk to the aquatic environment and identify appropriate decision criteria.

Due to the increasing intensity of debate on environmental hazard and risk assessment in Europe, the following additional terms of reference were subsequently added:

- critically review the proposed approaches to aquatic risk assessment;
- consider the practical implications for known environmental contaminants;
- recommend a scientifically sound approach for defining maximum levels of substances in the aquatic environment.

A first report is now presented. It has to be considered against legislative developments (see Section 4.1 below).

The terminology used in this report is that developed by the United Nations Conference on Environment and Development held in Rio de Janeiro in June 1992 and is also used in the EC "Risk Assessment Directive" (EEC, 1993). It differs from that which the EEC used in its workshop on risk assessment held in Ispra in October 1990 (EEC, 1990). Appendix A describes the two sets of terminology and defines other important terms used in this report.

SECTION 2. MODUS OPERANDI

2.1 CRITERIA FOR SELECTION OF DATA

In compiling the EAT database the following criteria were applied for the selection of data:

- data should be drawn from original scientific publications rather than from reviews or unpublished reports; reviews and databases were used to identify the source of original material;

- biological test methods employed were described, or reference was made to an appropriate published method;

- methods for the chemical analysis used to define the exposure concentrations of the test substance were described or referenced; thus all data represented measured rather than nominal concentrations;

- in the cases of limit tests, non-toxic substances and water solubility problems the following special criteria were applied:

  Results with substances that were non-toxic in limit tests were ignored and those which were non-toxic at 10,000 mg/l were rejected for calculation since they are considered environmentally irrelevant (they are included and indicated in the EAT database with 9999.9999). Potential problems regarding substances with low or very low solubility in water were not taken into account and the analytically verified values were taken as valid data. Studies in which vehicles (solvents, dispersants, etc.) were used to prepare sparingly water-soluble substances for testing were not excluded.

Due to the application of these selection criteria it is evident that for individual substances the database may not present the results of all valid studies performed. It is recognised that the data omitted may be of value towards other objectives than those of the present work, e.g. towards classification and labelling. On the other hand some of the data included may not be suitable for these purposes, particularly where in the case of sparingly water-soluble substances toxic concentrations could only be generated with the aid of vehicles.

2.2 SCOPE AND DEFINITIONS

Data were recorded for freshwater, estuarine and marine species.

It was found necessary to re-evaluate the definitions used for test duration, especially those used to describe 'acute' or 'chronic' toxicity. For example, a study of a few days represents a small fraction of the life cycle of a fish, a whole generation for a small crustacean or many generations of an alga.

To overcome the above inconsistencies the following definitions were applied:

> Acute exposure in animals covered any period up to one third of the time taken from 'birth' to sexual maturity provided that the animal could survive in good condition without feeding for such a period. Exposures were defined as sub-chronic if they were equivalent to no more than one third of the time taken to reach sexual maturity but feeding was required. Any more lengthy exposure was defined as chronic. For algae, chronic studies were taken to be those longer than 12 hours. Examples are given in Table 1.

Table 1 Definitions of Test Duration

Species	Usual maximum duration of test		
	Acute	Sub-chronic	Chronic
Algae	12 h	-	> acute
Daphnid	48 h	72 h	> subchronic
Zebrafish	96 h	30 d	> subchronic
Rainbow trout	7 d	250 d	> subchronic

Additional definitions may be found in Appendices A and C.

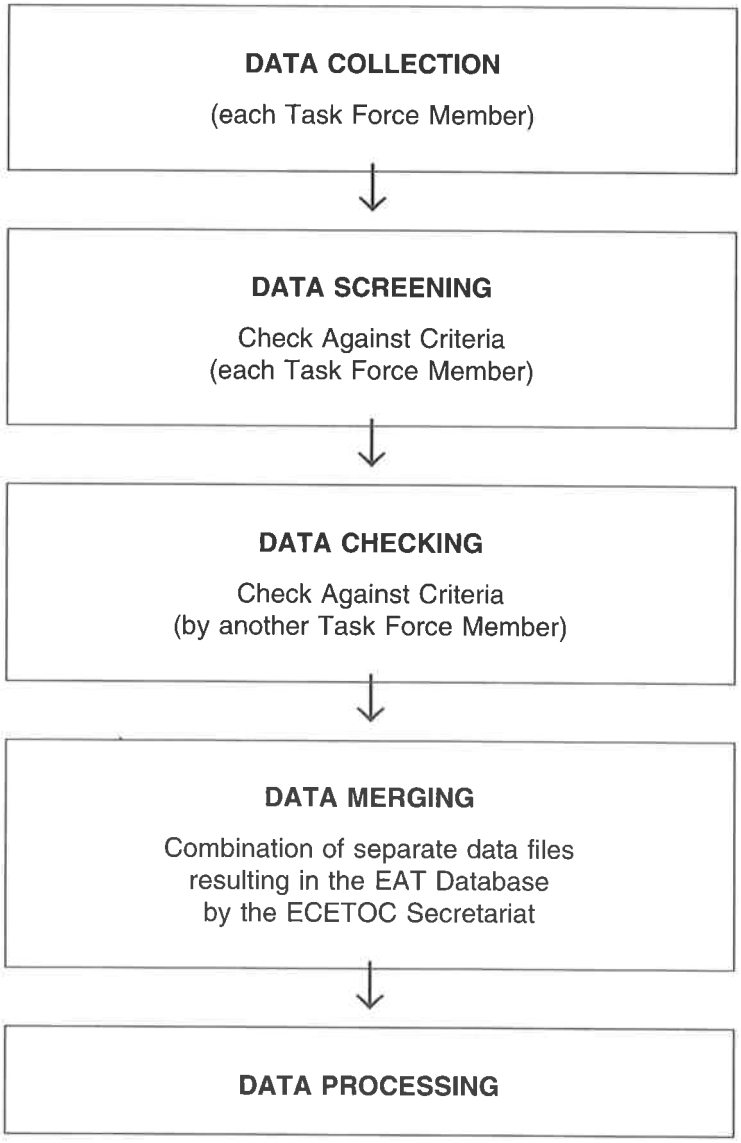
2.3 LITERATURE

Literature was gathered and screened according to the agreed criteria set out above. Applying the criteria, 530 papers were reviewed; of these 222 papers were found to be suitable for inclusion in the EAT database. The major reason for rejection of information was the failure to measure the toxicant concentration during the assay period. The literature screened dated from 1970 to 1991.

2.4 THE EAT DATABASE

A computer-based storage and retrieval system was established to aid in the processing and evaluation of the collected data which additionally permitted statistical analysis. A diagrammatic overview of the system used for data collection and evaluation is given in Figure 1. A more detailed description of the procedures used as well as a listing of the present data arranged according to Chemical Name and a CAS-No index are given in the separate Appendix C to this report.

Figure 1: Procedure of Data Collection and Evaluation



SECTION 3. RESULTS AND INTERPRETATION OF THE EAT DATABASE

3.1 MATHEMATICAL PROCEDURES FOR DATA EVALUATION

The evaluation of the data was carried out with a statistical software called SAS Version 6.04 for Personal Computers (SAS Institute Inc., Cary NC, USA) which is a menu-driven multi-window program. It offered the possibility of choosing data from the database by a selection procedure and of performing simple counts, frequency analyses, correlation studies, acute:chronic ratios and Hazen-distributions.

The principles and basic details of data evaluation in SAS are explained below.

3.1.1 Definition of "Value" and "Mean Value"

As used here, the term "value" represented an individual observation, such as an EC_{50}(96h) expressed in mg/l and was always related to one substance. Where relevant the toxicity data refer to the active form of the test material e.g. undissociated ammonia. Further detail is provided by Alabaster and Lloyd (1982).

Where for a certain selection (e.g. acute values for *Daphnia magna*) several values were found, the geometric mean of the values was calculated and this mean value used in the further evaluation.

3.1.2 Correlation Studies and Definition of "Data"

A major technique used was correlation analysis, such as species-species or acute-chronic relationships.

The values used to establish correlations were either mean values, if several values were available for a substance, or single values, if only one value was available. In the evaluation and correlation studies, the term "data" is used for both types of values.

The correlations were established using a special form of weighted linear regression for analyses written in the SAS-macro WFUNREL. Two characteristics of the data made this necessary:

- all data used to establish the correlations were (or were derived from) measured values (e.g. NOECs at the y-axis and LC_{50}s at the x-axis) with the consequence that both axes are subject to error;

- each point in the correlation was created by a single "x"-axis data point and a single "y"-axis data point; both data points, however, may be created from an unequal number of values (e.g. three values contributed to the "x"-axis data point and 15 values contributed to the "y"-axis data point).

WFUNREL took both characteristics into account by correcting for the error on both axes and including weighting factors for the different numbers of values which contributed to the "x"- and "y"-axis data point.

A typical printout is presented as an example (Figure 2). The size of the circle at each point reflects the number of "horizontal axis" and "vertical axis" values considered for each data point.

WFUNREL obtained the intercept and the slope parameters of a functional relationship and calculated the standard deviations and confidence intervals using an iterative procedure until the slope changed by less than 1%, which gave the final intercept and slope parameters.

3.1.3 Example: Calculation of Acute to Chronic Ratios

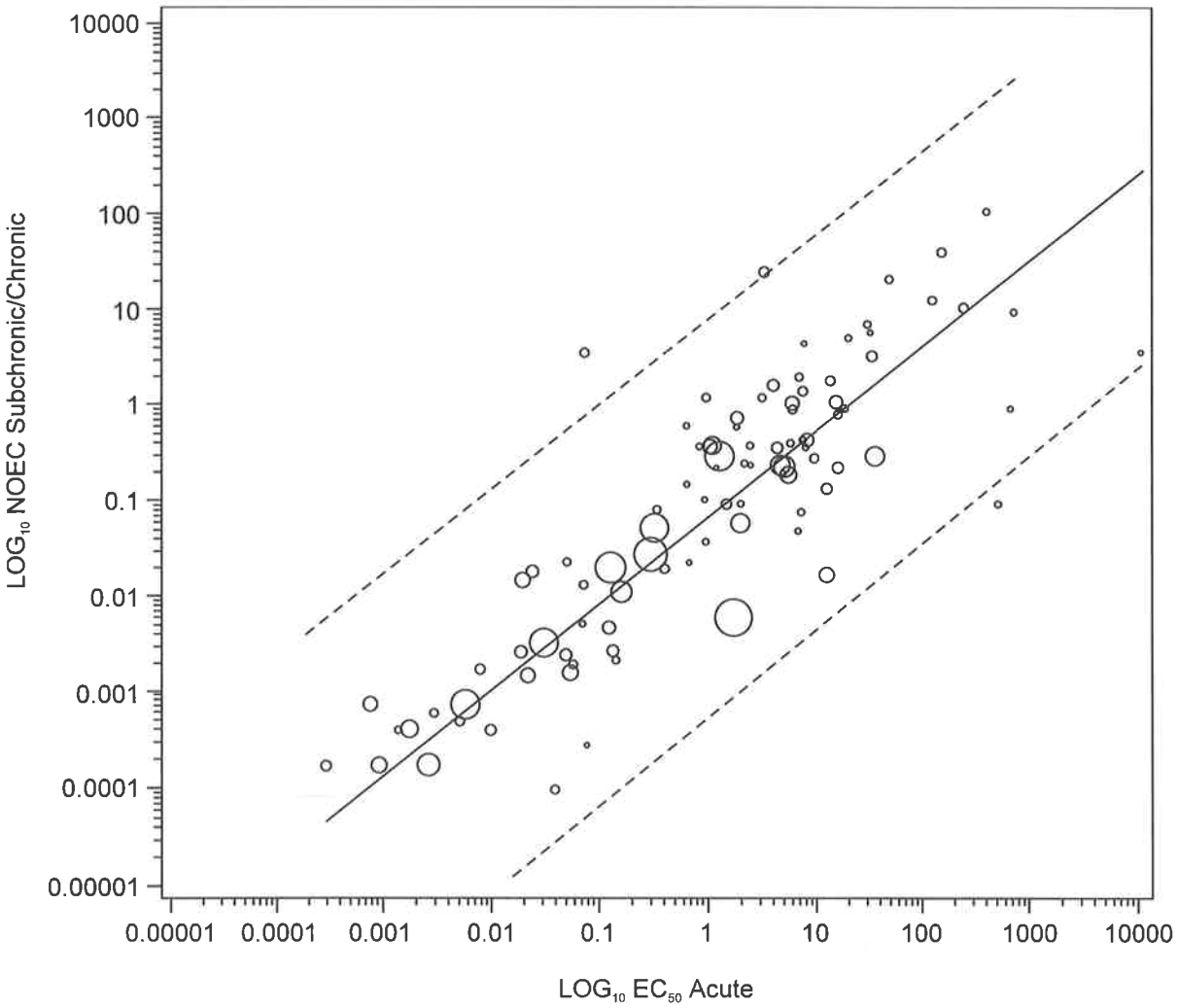
In relation to this calculation, "acute" data were defined as the acute EC_{50} and "chronic" data were defined as the chronic or sub-chronic NOEC. All other data (e.g. acute NOECs or chronic EC_{50}) were excluded from these calculations.

The data used to perform acute:chronic ratios were selected according to the following three methods:

- for each substance all available values were taken and geometric means were calculated for the combined species within "acute" or "chronic/subchronic" as defined in Table 1, this was the approach used earlier by ECETOC (1993);

- for each substance and each individual species geometric means were calculated for all the values for single species within "acute" or "chronic/subchronic" as defined in Table 1;

Figure 2 EAT Database: Mean Scores for Individual Chemicals



(96 points; F statistic (p-value) = 712.52 (0.000); R-squared = 0.883)

- the selections described before were refined in order to focus the acute:chronic ratios more precisely by choosing only data related to the following test durations:

  Daphnid acute: 48 hours,

  Daphnid chronic: 21 days,

  Fish acute: 96 hours,

  Fish chronic: 14-42 days.

3.1.4 Hazen Distribution

After calculating the ratios, they were ranked in ascending order and a "Hazen-percentile" was assigned to each.

The percentile is given for the n^{th} substance as described in Equation 1:

$$\frac{100}{2x} + \frac{100(n-1)}{x} \qquad (1)$$

where x is the total number of points in the analysis.

Thus, if there were ten points in the series the ones with the lowest (n=1) and highest (n=x) ratios were assigned, respectively:

$$\frac{100}{2\times10} + \frac{100(1-1)}{10} = \frac{100}{20} = \text{the 5 percentile} \qquad (2)$$

$$\frac{100}{2\times10} + \frac{100(10-1)}{10} = \frac{100}{20} + \frac{900}{10} = \text{the 95 percentile} \qquad (3)$$

This simply created a symmetrical plot for the cumulative distribution of the ratios as their percentiles, avoiding the statistical improbability of 0% or 100%.

3.2 SIMPLE COUNTS

Simple data point counting as presented in detail in Appendix D provides information on the nature and use of the substances included in the EAT database, the test organisms and the sensitivities between species for the various chemical classes of substances.

The sensitivity of the test organisms to the tested substances is evenly distributed on a log-normal scale (see Appendix D, Figs. D.6 and D.7). In differentiating between the sensitivity of the single groups of test organisms, it is obvious that bacteria tend to be the least sensitive whereas invertebrates (daphnids and non-daphnids) form the most sensitive component of the species spectrum tested (see Figs. D.8-D.13).

The distribution of the sensitivities (chronic NOEC) against the different chemical classes is in a log normal scale. However, particular groups can be distinguished. Heavy metals, aldehydes, organic

sulphur-, nitrogen-, and phosphorus compounds, esters and organometals are the more toxic groups whereas organic acids, alcohols, ketones and aromatic hydrocarbons are less toxic (see Figs. D.14-D.29).

3.3 INTERSPECIES COMPARISONS

Using the software described in Section 3.1, comparisons can be made between toxicity values for different species. For example, the capacity of data on one species to be used to protect all other species may be assessed. Various levels of specificity may be applied to this exercise (e.g. focusing on small ranges of exposure). As an example, the results of comparing the acute EC_{50} values of various "standard" species with the acute EC_{50} values of species or groups of species of the same or different taxonomic class are given in Table 2, Part A. It shows the percentage of substances where the standard species was more sensitive than "other" species (although the other species may include standard species). Table 2, Part B, shows the proportion of substances for which the interspecies acute EC_{50} ratios lay within a factor of 2.0. These approaches can be refined in subsequent work.

From the first assessment it can be provisionally concluded that rainbow trout (VF/SG) data would protect non-daphnid invertebrates for 90% of all substances and that data for algae (PA) would protect fish for 85% of substances. However, this is a preliminary assessment for illustrative purposes only and will be refined in subsequent reports. For interest the full data sets are given in Appendix E.

As well as comparing the relative sensitivity of a standard species to other (non-standard) species it is possible to use the regression analyses to describe the "goodness of fit" for each comparison. The proximity to 1.000 of the R-squared values shown in Table 2, Part A, demonstrates that the model provided a sufficient explanation for the distribution of the data.