Ann Ist Super Sanità 2002;38(2):119-129

Aquatic effects assessment: needs and tools

Silvia MARCHINI

Laboratorio di Tossicologia Comparata ed Ecotossicologia,
Istituto Superiore di Sanità, Rome, Italy

Summary. - In the assessment of the adverse effects pollutants can produce on exposed ecosystems, different approaches can be followed depending on the quality and quantity of information available, whose advantages and limits are discussed with reference to the aquatic compartment. When experimental data are lacking, a predictive approach can be pursued by making use of validated quantitative structure-activity relationships (QSARs), which provide reliable ecotoxicity estimates only if appropriate models are applied. The experimental approach is central to any environmental hazard assessment procedure, although many uncertainties underlying the extrapolation from a limited set of single species laboratory data to the complexity of the ecosystem (e.g., the limitations of common summary statistics, the variability of species sensitivity, the need to consider alterations at higher level of integration) make the task difficult. When adequate toxicity information are available, the statistical extrapolation approach can be used to predict environmental compatible concentrations.

Key words: aquatic toxicity assessment, QSAR, toxicity testing, data analysis, extrapolation.

Riassunto *(La valutazione degli effetti sull'ambiente acquatico: necessità e strumenti).* - Nella valutazione degli effetti che le sostanze inquinanti possono produrre sugli ecosistemi esposti si possono seguire diversi approcci, a seconda della qualità e quantità di informazioni disponibili, i cui vantaggi e limiti sono discussi facendo riferimento al comparto acquatico. In assenza di dati sperimentali si può adottare l'approccio predittivo utilizzando relazioni stuttura-attività validate, ma solo l'applicazione del modello appropriato fornirà stime di tossicità affidabili. L'approccio sperimentale è fondamentale in ogni procedura di valutazione di pericolosità per l'ambiente, anche se l'estrapolazione da un limitato set di dati di laboratorio su specie singole alla complessità dell'ecosistema presenta molti elementi d'incertezza, tra cui le limitazioni dei comuni endpoint statistici, la variabilità della sensibilità delle specie, le alterazioni a livelli di integrazione più elevati. Qualora si disponga di informazioni tossicologiche sufficienti ed adeguate, l'approccio statistico di estrapolazione può fornire valori di concentrazione compatibili con l'ambiente.

Parole chiave: valutazione di tossicità acquatica, relazioni struttura-attività (QSAR), test di tossicità, analisi dei dati, estrapolazione.

Introduction

Purpose of environmental effects assessment process is to provide estimates of the concentration of a chemical (PNEC, predicted no effect concentration) below which unacceptable effects on the ecosystem are not likely to occur. It represents an essential part of the complex procedure of environmental risk assessment of chemical substances [1] and it provides the scientific rationale, beside the economical, technological, social considerations to fix legal limits (WQO, water quality objectives) as well [2].

The process of effects assessment implies the identification of the intrinsic ability of a substance to produce adverse effects and their quantification. To this purpose, focussing on the aquatic compartment, the toxicity data bases are evaluated considering the information available on: i) physico-chemical characteristics of the compound, which represent the basic knowledge which permits to identify the possible routes of exposure, the potential to accumulate into the organism following bioconcentration from direct uptake from water and biomagnification through the food chain, the potential to adsorb onto sediments, and to persist; ii) quantity and quality of toxicity data. A critical evaluation of the adequacy of data is of the utmost importance, in fact only reliable data (obtained from a well described test conducted according to standard methods) relative to ecologically relevant endpoints (e.g. appropriate endpoints from tests of appropriate duration) should be retained for the assessment; iii) biological and ecotoxicological monitoring studies, which provide additional information on realistic scenarios, following practical exposure.

Indirizzo per la corrispondenza (Address for correspondence): Silvia Marchini, Laboratorio di Tossicologia Comparata ed Ecotossicologia, Istituto Superiore di Sanità, V.le Regina Elena, 299, 00161 Roma. E-mail: marchini@iss.it.

Xenobiotics may affect ecosystems in various ways and at various extent, leading to changes in their structure and function, hence effects assessment should be made with reference to both these alterations.

In practice the amount of information available for a substance may vary tremendously, from just physico-chemical data to experimental results from field studies. Accordingly, depending on the quality and quantity of toxicological data, different strategies can be followed to pursue effects assessment, which permit an optimal use of the available data:

- when experimental data are lacking we can follow a predictive approach applying QSAR (quantitative structure-activity relationship) models which provide estimates of the toxicity;
- when the data set includes only few acute and/or chronic data, a pragmatic approach is usually followed, which consists in applying assessment factors to the lowest toxicity value within a single endpoint;
- if several data are available on a certain toxicity endpoint, extrapolation methods can be applied, which provide toxicity estimates for the (hypothetical) most sensitive species in the ecosystem or estimates of the maximum tolerable environmental concentration (MTC);
- a closer approximation to reality is achieved with microcosm, mesocosm up to field tests, which permit a refined level of assessment.

To this increasing bulk of information corresponds an increasing level of confidence in the assessment, which is reflected by the decreasing size of application factors used in the extrapolation from experimental laboratory data to natural environment.

In the following, the advantages and limits of the different approaches to environmental effects assessment and some key issues relevant to the statistical analysis of toxic effects and data interpretation are discussed with reference to the aquatic compartment, along with some examples.

Predictive approach

Often the toxicity assessment of environmental pollutants on aquatic organisms is difficult because of the lack of experimental data, even the most simple ones. QSARs are mathematical models which quantitatively describe a biological response to a chemical in terms of structural parameters. In absence of experimental data, they permit a very preliminary assessment of toxic effects of chemicals and can be used for setting priorities among them.

It is recognized that the QSAR approach is successful in predicting toxicity only if the proper model is used. Since QSAR models are valid for chemicals with the same mode of action, it is crucial to assign a chemical to the proper mechanistic class.

For environmental toxicants four broad classes of mode of action have been identified [3-5]:

- class I: non-polar narcotic chemicals. They possess the minimum (baseline) toxicity which can be predicted by their hydrophobic properties only (log P);
- class II: polar narcotics. They are less inert chemicals, whose toxicity is slightly higher than that predicted by the baseline equation of class I chemicals. This excess toxicity (toxic ratio) is expected to be up to 10;
- class III: unspecifically reactive chemicals. They can be 10-10 0000 times more toxic than predicted by class I equation and their modelling requires a reactivity parameter in addition to log P;
- class IV: specifically acting chemicals, which are receptor-mediated toxicants.

For non-polar chemicals, a number of reliable baseline equations, are available for different organisms and endpoints, and well established QSARs have also been calculated for acute toxicity of polar compounds to fish and Daphnia [5-7]. For both classes of narcotics the variations in toxicity is satisfactorily explained by their lipophilicity alone and log P is the only descriptor. A multitude of models have been developed for reactive chemicals but their use has not been recommended because not sufficiently validated yet.

Classification schemes have been produced which define a provisional set of structural rules that can assist in class assignment [3, 4, 8-10]. With the purpose of validating these classification schemes and criteria, we conducted a study to investigate the toxicity of halogenated benzenes and toluenes to *Daphnia magna* [11]. According to structural rules, halo-benzenes and - toluenes belong to either class I (compounds with halogen on the ring; iodinated chemicals excluded because potent alkylating agents) or III (toluenes containing a good leaving group at the benzylic carbon). This classification is generally supported by experimental studies mainly carried out with fish and on chlorinated compounds only. Since mode of action depends on species, compound, and toxic endpoint, it was considered important to validate the classification in other taxa and for other halides (29 compounds with a variety of halogen substitution on the ring and/or methyl group). Main results are shown in Fig.1, where toxicity (LC50) are plotted against log P. The model obtained with all tested compounds (excluding outliers) (Eq. 5) does not differ much from the model with only ring-chlorinated compounds (Eq. 2); on the contrary both equations show lower slopes and higher intercepts than the general baseline model from literature (Eq. 1 [5]).

This indicates that halobenzenes and toluenes, which are located in the higher portion of the log P range, may behave somehow differently from other class I chemicals. They do have a common membrane partitioning process but they appear to possess also additional interactions with target site.

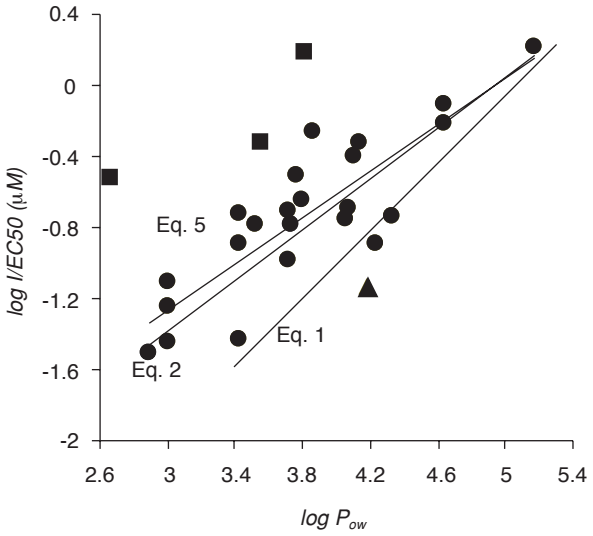

Fig. 1. - Experimental toxicity of aryl and benzylhalides to *Daphnia magna* and log P-dependent QSAR equations. Eq. 1 (non-polar narcotics, OECD, 1995): log 1/EC50 = 0.95 logP - 4.8 (n =17 [2],r = 0.99, s = 0.21); Eq. 2 (arylchlorides): log 1/EC50 = 0.71 logP - 3.53 (n = 7, r^2 = 0.86, s = 0.23); Eq. 5 (aryl- and benzyl-halides): log 1/EC50 = 0.65 logP - 3.2 (n = 23 [2]; r = 0.74, s = 0.23). (Please note that equation numbers refer to the numbering in the original paper).
(Reprinted with kind permission from Marchini *et al.,* 1999, *Environmental Toxicology and Chemistry* [11]). Copyright Society of Environmental Toxicology and Chemistry (SETAC), Pensacola, FL.

Based on the calculated toxic ratios between predicted (by Eq. 5) and observed toxicity, halobenzenes and toluenes were classified as follows (Fig. 2):

- ring-halogenated compounds are assigned to class I regardless the type and pattern of halogenation at the ring (iodine compounds included). They are in general very well predicted;

- the classification of benzylhalides is more complex. Chemicals act by different modes of action, depending on the type and number of halogen at the benzylic carbon. When only one H atom is substituted (by Cl or Br) they are more toxic than predicted and, like in fish studies, can be assigned to class III. When two Cl atoms are present, the toxicity is perfectly predicted by the baseline model. When the benzylic C is fully substituted, the toxicity will strongly depend on the type of halogen. In the case of F, the compounds behave like class I narcotics, while if the halogen is a good leaving group (like Cl) chemicals are likely to be very reactive (class III).

The QSAR models obtained in this study indicate that most halo-benzenes and -toluenes are actually baseline chemicals with some extra reactivity and as such form a subgroup within class I chemicals. These results suggest the hypothesis, supported by findings from mechanistic studies such joint toxicity, toxic syndromes and critical body burden, that there are several different ways to produce baseline toxicity. Generally the classification scheme derived from fish studies has been confirmed; however, major deviations from structural criteria were observed for ring-iodinated compounds and aa-dichlorotoluenes that are not reactive, but behave as non-polar non-specific toxicants (class I chemicals). Obviously, these findings are valid for the test species and endpoint considered; expanding the investigation to other organisms and type of test would possibly lead to more complicated toxicity profiles.

Experimental approach

The ultimate goal of the effect assessment process is the protection of structure and function of the ecosystem, which can be defined as "a stationary entity of biotic and abiotic compartments, in which energy, material and information reside and flow in various forms" [12]. In Fig. 3 the ecosystem structure is illustrated at trophic levels by food web components and its function is described in terms of uptake and release of energy and materials and their transfer between levels. In a stressed ecosystem, changes in structure are reflected by altered populations and community composition and changes in function correspond to variation in productivity and in rates of uptake and release of minerals and gases.

The ecotoxicological testing for the evaluation of the chemical hazard is central to any risk assessment procedure and environmental regulation and management. Pollutants can affect biota at different levels of biological organisation and the experimental approach includes ecotoxicological testing of increasing complexity, ranging from screening tests of short duration (minutes, hours) at cellular or whole organism level, to full tests which span the entire life cycle of a single test organism (months, years) to field trials which deal with the population and community levels.

Single species tests, due to the possibility of high standardisation and control of experimental variables, allow to determine a direct causative link between exposure to chemical(s) and biological effect and to investigate the mechanism through which they elicit the toxic action. Both short- and long-term tests are necessary. The first can serve in the screening phase to identify the toxic potential of as many as possible chemicals to as many as possible organisms. The long-term tests permit a more refined hazard assessment as they reveal sublethal, reproductive, and delayed effects of chemicals. The investigation of substances with specific mode of action may require to be addressed with *ad hoc* tailored tests.

In an extensive effect evaluation, representative species of all the trophic levels present in the target ecosystem should be included. Consequently, an

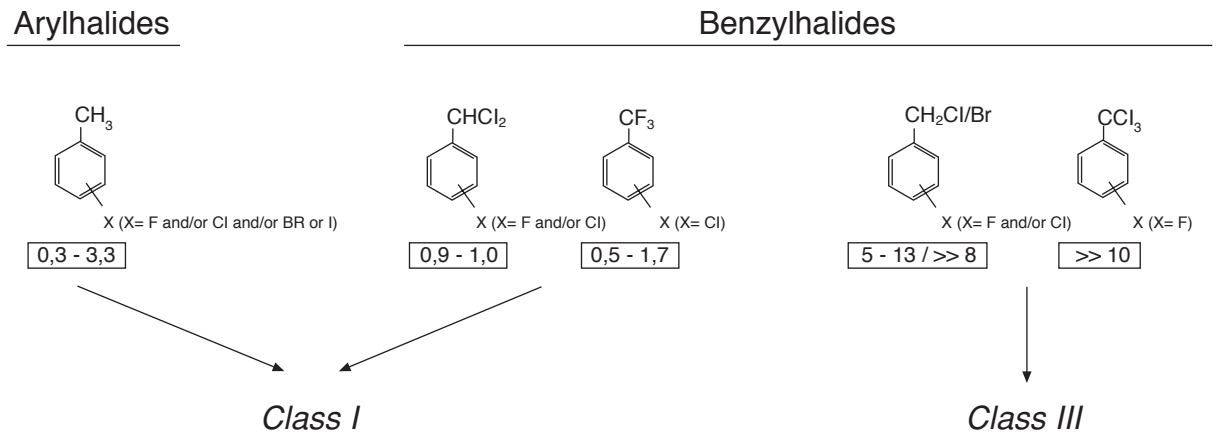

Fig. 2. - Classification of tested halogenated benzenes and toluenes by mode of action. Class I: non-polar narcotic chemicals; class III: unspecifically reactive chemicals. Toxic ratios (Eq.1 predicted EC50 divided by experimental EC50) are reported within squares.

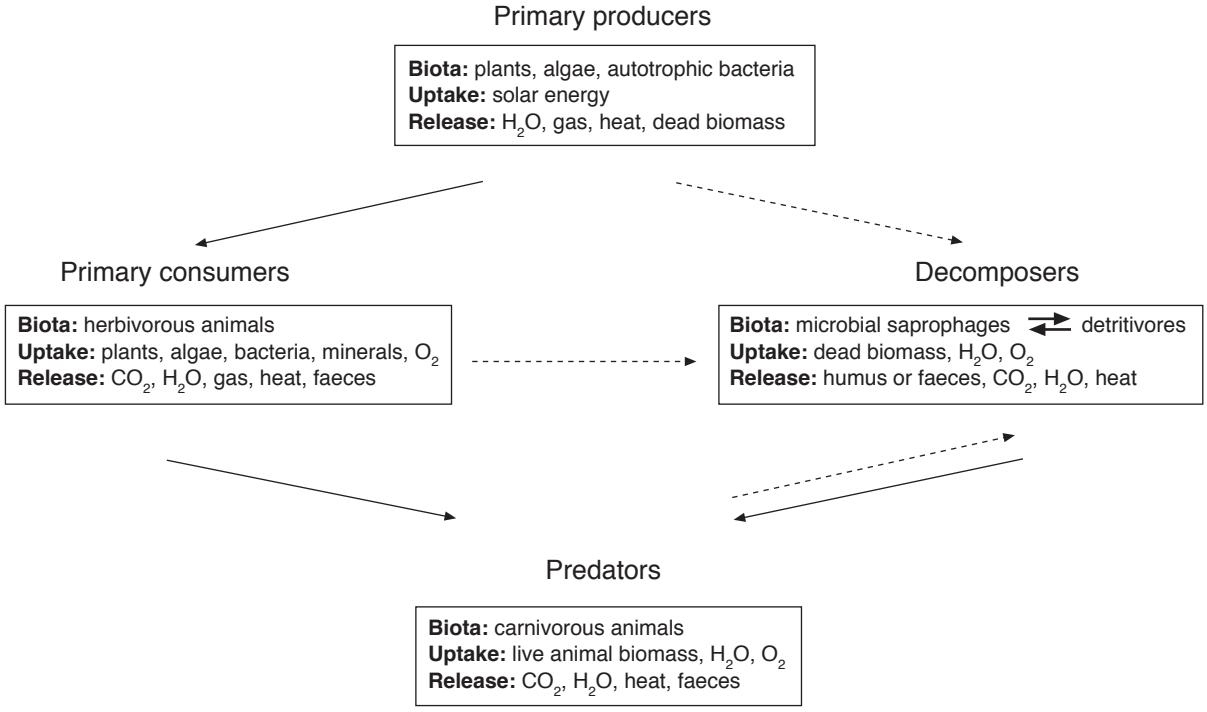

Fig. 3. - Schematic representation of biota structure, as trophic levels, in the ecosystem. Arrows indicate fluxes of energy and material among levels.

increasing effort is being world-wide devoted to the development and validation of standard methodologies, to increase the reliability and reproducibility of toxicity tests. In the frame of the OECD Testing Guidelines Programme the need for developing new guidelines has been identified within the general scheme in Fig. 4 and priorities have been indicated which are meant to fill major gaps in key taxonomic groups at each trophic level. Drafts are in the process of being finalized for sediment test with chironomids and plant test with Lemna.

Nevertheless it is accepted that the web of inter-relations within natural communities, and between these and the abiotic environment, cannot be described by the simple addition of effects toward single species, so that single species testing approach provides thresholds which inevitably allow a limited interpretation of what occurs at higher levels of organization. At the ecosystem level, the impact on biota should in fact be assessed for populations and communities (measuring variables such as population size and density, diversity and species

richness, frequency of species distribution), together with the changes of functional properties (measuring variables such as productivity, respiration, nutrient cycling) which reflect changes in flows of energy, material and (chemical, genetic, structural) information [12]. Furthermore, as the biota is influenced by the elements of the abiotic compartment, and *viceversa*, in a comprehensive risk analysis, alterations of abiotic components (air, sediment and water) should also be taken into account. From the kind and extent of modifications in all the above components will depend the ability of the ecosystem to recover from a stress.

Structural and functional variables can be monitored in field tests and, to a certain extent, in microcosm and mesocosm tests. These tests aim to simulate the ecosystem complexity, reproducing it at (different) smaller scales; they therefore represent the highest tier in the hazard assessment strategy. On the other hand, the higher the complexity of test system the lower the degree of standardization and reproducibility that can be attained. This fact, together with the great variability (in space and time) of natural ecosystems, partly limits the predictivity of such experiments, which can only be of probabilistic nature.

The statistical analysis of test results

A critical aspect of toxicity testing is the choice of the statistical method of data analysis. The great majority of data available derives from acute tests on single species, due to their obvious practical advantages. Acute, short-term tests are traditionally analysed using a static regression approach, in which a model is fitted to experimental data recorded at a certain time (e.g. 96 h) and it is used to estimate the respective EC/LC50. Since long, the limits of this toxicity descriptor have been recognised and it has been suggested that the usefulness of short-term tests could be improved making them more informative through the analysis of the shape and slope of the EC50 *versus* time curve [13].

Information on how effects build up over time can assist in the interpretation and evaluation of standard toxicity results and therefore are important for risk assessment. The way an organism responds to chemical exposure reflects the interactions among exposure, toxicokinetic and toxicodynamic factors, hence the knowledge of the mode of action of chemicals is important in relation to how a standard toxicity endpoint (e.g. fish 96 h LC50) should be interpreted. For aquatic toxicity, two different models of LC50 *versus* time can be taken into account, whose theoretical bases have been recently discussed by Verhaar *et al*. [14].

For narcotic chemicals that act through a non-specific mechanism, that is physical interaction with cell membranes, the interaction with target is assumed to be instantaneous (the binding is much faster than the uptake) and reversible. The critical body residues (CBR, i.e. the internal concentration at the time of occurrence of the effect) is constant in time and also among different compounds and different species [15]. This corresponds to a pharmacological model where the (lethal) dose is described by peak concentration at the target site (internal concentration) [16].

For reactive chemicals (that either react unspecifically with biological membranes or specifically with receptors) the interaction with target is assumed instantaneous and irreversible; for this pharmacological model the (lethal) dose is better expressed in terms of area under curve (AUC), that is the integral over time of the internal concentration [17]. It can be shown that the critical AUC (CAUC) is not expected to be constant among toxicants, because it will depend on the reactivity of each compound, and that the CBR will be time-dependent).

Accordingly, for the narcotic chemicals the time-effect relationship will follow an exponential function (which is determined by bioconcentration kinetics, in fact it is practically the inverse of the uptake curve) (Fig. 4). We can say that at steady state, when the bioconcentration equilibrium is attained, the LC50 becomes constant. For reactive chemicals an alternative model has been investigated, where the LC50 results to be time dependent (a function of 1/t) [14]. For this class of toxicants the time-effect relationship is described by an hyperbolic function (which is determined by cumulative inhibition of the receptor) and the bioconcentration steady state is often reached before the LC50 becomes

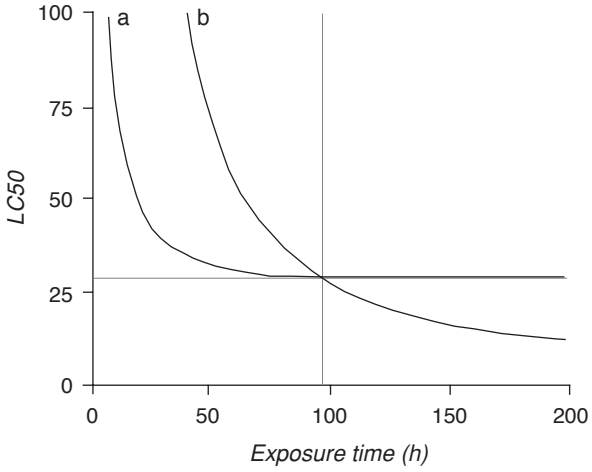

Fig. 4. - LC50 *versus* time models.
a) narcotic chemicals: LC50 (t)= LC50 (∞) /1-e $^{-k_2 t}$
b) reactive chemicals: LC50 (t)= CAUC/BCF x 1/ [t-(1-e$^{-k_2 t}$) /k_2]+ EC50 (∞);
CAUC = critical area under curve; BCF = bioconcentration factor, k_2 = first-order elimination rate constant. (Modified with kind permission from [14]).

constant. A major implication of the existence of these two models is, for example, that the acute toxicity of chemicals acting through a non specific mode of action, would be underestimated if only LC50 at a fixed time (e.g. 96 h) is considered.

Researches in such directions should be encouraged as a deeper knowledge of the mechanisms of toxic action would provide important tools for risk assessment (either if experimental or predicted toxicity data are available) and would contribute to design more efficient toxicity tests.

Conventionally, the results of (semi)chronic ecotoxicity testing have been expressed and reported in terms of NOEC, i.e. the highest concentration at which the measured response is not significantly different from the control, but its use as summary statistic has been largely criticized. The determination of the NOEC is carried out in two steps. First a parametric ANOVA is performed to check if the underlying assumptions regarding residuals distribution hold, then a multiple comparison procedure follows. The NOEC is widely used as a basis for risk assessment, but there is general consensus that it is not a satisfactory parameter to describe the level of no-effect. Major drawbacks associated with the use of NOEC are agreed to be:

- the NOEC can only be one of the concentrations selected by the experimenter;
- it is not possible to associate any precision to the estimate;
- we might not be able to detect any NOEC from our test and waste important information associated to the concentration effect curve;
- if the variability in the experiment is high the sensitivity of the statistical analysis will be low and only large differences can be detected [18], i.e. NOEC can correspond to large effects. NOEC has no correlation with NEC (no effect concentration) and therefore cannot be considered a "safe" concentration. For example, the analyses of the results of a reproduction ring test with *Daphnia magna* and of a ring test on fish growth showed that effects as large as 37-38% respectively were found to correspond to the NOEC values of a reference toxicant [19, 20]. We obtained similar findings in a short-term test with fish larvae exposed to benzene derivatives [21]. The use of a statistical test with unknown power can lead to the paradox of the NOEC being equal or larger than the LC50 [21, 22];
- generally the magnitude of effect at the NOEC is not reported in literature, and this can be misleading for effect assessment.

Recommendations have been given that the statistical approach to aquatic toxicity data analysis move away from NOEC calculation and shift toward point estimates based on regression models (ECx), which possibly incorporate the time of exposure [23]. Anyhow, the analysis by regression is not without problems. Major concerns arise from the choice of the model, e.g. log logistic, log probit. In fact, when we move from the LC50, which is a robust estimate, to a small effect level, e.g. LC5, the estimated concentration will strongly depend of the model chosen (x) in the EC_x estimate.

Since the traditional summary statistics (EC50/LC50 and NOEC) vary with time of exposure depending on the properties of both chemicals and organism (see also next paragraph) and considering the problems discussed above, it has been argued that the description of toxic effects by standard model does not guarantee the comparability of results from standard tests. To overcome these limits, an alternative mechanistic model (dynamic energy budget, DEB) has been proposed, which is based on fluxes of energy through an animal (related to the various physiological processes) and their variation during different life stages (see [24], and references thereof). Very briefly, the DEB approach describes long-term toxic effects by means of two parameters (NEC and a tolerance concentration) using information from three components: the kinetics, the effects, and the physiology of the species. This model, because based on the physiology of test organism, permits a straightforward and biological sound interpretation of toxic effects, also in terms of population dynamics. Mechanistic models have the advantage to produce reliable and informative thresholds of toxicity, and may represent valid alternatives to empirical models, provided that they fit the data as good as the latter ones.

The problem of the approach to data analysis, with special reference to available standard guidelines for aquatic toxicity testing, has been focussed and discussed during an *ad hoc* OECD workshop [23]. A dynamic regression approach to data analysis, which provide time-dependent models, has been proposed, for both short- and long-term test, and its adoption has been put forward by the scientific community. In these models the biological response is a function of both concentration and time, so that EC estimates can be calculated at several times and time to response can be estimated as a function of concentration. The inclusion in data analysis of the time variable would require minor modifications of tests design. Work is in progress to compare the adequacy of different types of dynamic models (both mechanistic and empirical) and identify, for the various tests and endpoints, which is the optimal test design and the most appropriate percentage of effect to consider.

Extrapolation to ecosystem

A common feature of the effects assessment procedures, adopted in environmental regulation and management of different countries and agencies in the world, is to rely upon a small battery of toxicity tests focussing on organisms from different trophic levels to

predict threshold level of sublethal effects in the ecosystem, both for single chemicals and complex mixtures. In general, only toxicity data on pelagic organisms, exposed via the water phase, are available, and this may not be adequate to assess the effects to bentonic organisms whose uptake occurs also via the sediment phase. The inclusion in the battery of a test with sediment is therefore highly recommendable, especially for substances with a high potential to sorb onto sediment, to accumulate and persist. Usually, the battery consists, as a minimum, of an alga, a crustacean (Daphnia), and a fish, and its adequacy has been empirically demonstrated. Studies have been conducted to search for the "best" battery; for example, the analysis of a large set of sub(acute) toxicity data showed that the sensitivity of the standard test battery could be improved by the addition of a diatom [25]. A more refined assessment is possible by means of a specific environment-oriented testing strategy that takes into account the species potentially exposed.

In extrapolating from few lab data to the ecosystem we make two important assumptions, that are at the basis of any ecological effects assessment; the first is that the ecosystem is protected when the most sensitive species is protected, the second is that the function of ecosystem is protected when its structure is protected. Considering the complexity of the aquatic environments and the multiple processes and interactions present in the ecosystems (see previous chapter), it is certainly an over-simplification to base the effect (and risk) assessment on a bunch of short-term laboratory toxicity data obtained on single species, nevertheless this is, unavoidably, the most frequent occurrence. In a preliminary step, it might be even necessary to provide a rough estimate of the concentration "without" effect based on only one LC50 or even on predicted toxicity values.

The elements which contribute to the uncertainty in the extrapolation of a "safe" concentration for the natural environment are multiple (the list could be longer):

- few standard species to natural community extrapolation;
- laboratory conditions to natural environment extrapolation;
- fresh-water to marine organisms extrapolation;
- short- to long-term effects extrapolation;
- pelagic to benthonic organisms extrapolation;
- data variability;
- intra- and inter-species variability;
- different sensitivities of life stages;
- different routes of exposure;
- single compounds to mixtures extrapolation;
- NOEC *vs* NEC;
- 48/96 h LC50 *vs* stationary LC50;
- indirect effects (e.g. via the food chain);
- interactions among species;
- interactions between biotic and abiotic environments;
- recovery rates.

To take all these elements into account, arbitrary, empirical assessment factors are applied to the lowest toxicity values (of the most sensitive endpoint of the most sensitive species) to derive a PNEC or WQO for environmental (industrial) xenobiotics.

The size of the application factor will depend on the chemical properties (e.g., persistent, bioaccumulative) and on the confidence we have in the available data: the higher the quality and quantity of information (measured/predicted values, number of trophic levels, taxonomic groups and feeding strategies, test duration, type and relevance of endpoint), the higher the confidence and the smaller the factor. Traditionally, assessment factors are raised with a factor of 10 (the highest factor is applied to acute LC50 and lowest to mesocosm/field test results). However, the factor size is also modulated in consideration of the number of available NOEC from long-term tests and the fact that they derive or not from the species/trophic level resulted most sensitive in acute tests. At any rate, it should be underlined that final assessment should always rely upon expert judgement, which is the main tool for a critical weighing of the information. This is especially true when old data from tests not conducted according to standard methods are to be evaluated. As the available information can be very variable, some flexibility should be allowed to experts in choosing the assessment factor they believe more appropriate. Guidance to the use and choice of appropriate application factors can be found, among others, in European Commission [1], CSTE/EEC [2], OECD [26] and Zeeman [27] for US EPA.

Each of the uncertainty elements in the above list represents a key issue in ecotoxicological research, from which a better knowledge of the toxicological properties of xenobiotics and more refined tools for reducing uncertainty and for a more sound risk assessment are expected. This would also provide the rationale for a sound choice and application of less empirical and more scientific assessment factors.

Since the toxic concentration range of a single chemical among species can span several orders of magnitude, variation in sensitivity represents a major source of uncertainty in determining the effective environmental concentration. The search for a sensitivity pattern among species to find the "most sensitive" one, has therefore been the objective of many comparative analyses, where the toxicity of organic and inorganic pollutants, with a variety of mode of action towards many aquatic species belonging to various taxonomic groups, have been analyzed [25, 28-31]. These studies all showed the lack of a susceptibility order among species and to date a general theory for the relative susceptibility does not exist. This is due to the fact that the susceptibility varies in relation to the mode(s) of action of the single substance and also depends on the morphology, behavior

and physiology of the organism (especially important are differences in the rates of uptake, metabolism and elimination of toxicant). Limited to non-specifically acting chemicals, fat content can account for different sensitivity, and kinetics-related factors such as body size and surface area have been individuated as elements influencing toxicity [31]. A general observation is that, as expected, chemicals with a non specific mode of action show small variation in toxicity towards different organisms in comparison with specifically acting toxicants which are organism-selective. This is easily explained with the fact that the unspecific partitioning processes across membranes (which represent the rate limiting factor for narcotic compounds) are common among different organisms while the specific interactions vary depending on the number and type of receptors present in each organism. For specifically-acting chemicals, the "most sensitive species" will therefore be that possessing the target biomolecules mediating the specific effect [7, 32].

In a study conducted with several test fish species recommended by OECD [33], toxicity differences among species have been reported, which have been explained with the fact that the enzymes involved in the metabolism of xenobiotics (like monooxigenases and glutation-transferase) can vary up to 10 times. Interesting is that the variation in susceptibility between and within lower taxa was found similar, up to four orders of magnitude [28]; anyhow, in comparison with invertebrates and algae, fish show a relatively narrow range of sensitivity [29, 34]. Worth of consideration is the finding that for some chemicals differences among species within the same taxonomic group may be greater than those among organisms from different taxa. For example, pentachlorophenol (which is an uncoupler of the ossidative phosphorilation) was 2 log unit less acutely toxic to a bentonic crustacean than to Daphnia [29], and benzophenone showed a difference of 2 log units in acute toxicity to *Daphnia magna* and *Ceriodaphnia dubia* [35], while in both studies the toxicity to Daphnia compared well with that of fish.

Since, at best, only a limited number of data can be collected, a crucial point in effects assessment is therefore the choice of species/taxa (and endpoints) to be included in the test battery. This is especially true if a statistical approach to extrapolation, which implies the analysis of sensitivity distribution of the species in the community, is followed. Probability distributions of species sensitivities describe variation among species in response to toxicants and can be used to extrapolate the toxic concentration for the most sensitive species [36, 37]. In the extrapolation approach, contrary to what occurs when the approach with application factors is followed, all available toxicity data are used. Extrapolation methods are based on the assumption that both the measured LC50/NOEC of a limited set of tested species and the unknown LC50/NOEC of all species in the natural community can be regarded as random trials from the same frequency distribution; in other words both the sensitivity and the sensitivity distribution curve of laboratory and field species are similar.

Various extrapolation methods have been proposed [36, 38, 39], which differ in the type of assumed distribution (log-triangular, log-normal, log-logistic, respectively) and in the requirement of type and number of toxicity data (chronic NOEC from ≥ 8 families, or from ≥ 4 representative species or from ≥ 4 sensitive species from different taxa, respectively). Common purpose of these methods is the estimate of a concentration (final chronic value, FChV, or hazardous concentration, HC) at which a chosen x% (the most sensitive fraction) of genera/species in the community is at risk. In practice, x is conventionally posed equal to 5, so that the ChV or HC would protect 95% of genera/species. The rationale behind this, is the assumption that redundancy in the ecosystem is such to permit some loss in the communities without significantly altering its structure and function [40]. To take into account the uncertainty in the extrapolated value when few data are used, the lower tolerance confidence limit of the 5^{th} percentile is used in ecological hazard assessment.

Against the assumptions underlying statistical extrapolations it can be argued, in addition to the bias related to the above listed uncertainties, that, within the lost 5% of species, a critical species might be comprised [41] and any loss can alter the structure and function of the community [42]. Furthermore, it should be observed that the above mentioned extrapolation models do not incorporate factors at level higher than organism, such as population dynamics and interactions among species, so that their use for an integrated effect prediction at ecosystem level is hampered. For several chemicals, assuming a log-logistic distribution of LC50 values, it has been found that the factors obtained dividing the average chronic EC50 by the HC for sensitive species were much higher than standard assessment factors [37]. This means that the application of commonly used "safety" factors is likely to be not protective for the community. It has also been shown that, for a reliable environmental effect assessment, the knowledge of the species potentially exposed is of decisive importance. In fact, on the assumption about community diversity depends the variability in toxicity (width of distribution curve) and hence, the estimated toxic level for the hypothetical most sensitive species. This is illustrated, for example, by an exercise with phenol by Nendza *et al.* [43], who calculated a significantly decreasing of the toxic threshold if different local scenarios (few or many fish species and/or invertebrates) were considered.

The pros and cons of using one or the other extrapolation method have been explored [44]. The conclusions of an OECD comparison exercise were that

the three above mentioned methods give comparable ChV (95%) estimates [26]. Nevertheless, the problem of checking the validity of assumptions (e.g. the distribution model) remains. A recent analysis, that compares the fit of different models to many data sets, illustrates the importance of selecting the appropriate sensitivity distribution and shows that, on many occasions, the use of a log-normal distribution may not be satisfactory [45]. To overcome this problem, the use of a bootstrap technique has been proposed, which is distribution-dependent and, moreover, permits to calculate the number of species necessary to accurately estimate the HC.

To date, statistical extrapolation methods are used in the USA and some European countries (e.g. the Netherlands) to fix WQO, but their use in risk assessment of new and existing chemicals is suggested only in support of the approach with application factors, because judged not sufficiently validated yet [1].

Conclusions

The complicated task of effects assessment can benefit from the predictive as well as the experimental approach. Validated QSARs represent useful tools which can be used in risk assessment to set priorities among risk chemicals and, define the strategy of further studies, and, in conjunction with experimental data, can help in understanding the toxicological properties of chemicals. Nevertheless, QSAR models can provide accurate toxicity estimates provided that appropriate models are applied. Further (QSARs and mechanistic) studies are necessary to better identify the structural moieties responsible for different modes of toxic action in various test systems. This would help in assigning a chemical to the right mechanistic class and make QSARs more powerful predictive tools in hazard assessment, both in absence and in support of experimental data.

Testing at both single species and complex system levels contribute to understand and quantitatively estimate the impact of xenobiotics on the environment.

Testing systems at a high level of integration, permits a more reliable effects evaluation while single species testing allows to clarify direct causative relations between exposure and biological responses. For practical reasons, the latter remains the basis of the hazard assessment procedure; it is therefore highly desirable to improve the design of standard tests, both short- and long-term, so that as much information as possible can be extracted. Current summary statistics used in the analysis of data from standard tests are not satisfactory toxicity endpoints for regulatory purposes. To use toxicity data more efficiently and reduce endpoints ambiguity, point estimates can be calculated from dynamic (i.e. time dependent) models. Where possible, the application of mechanistic models, which integrate energetics, would provide more biologically relevant thresholds of toxic effects, also beyond the species level.

In the extrapolation from experimental data to the real environment, both the approach with application factors and the statistical one have evident limits. What is of utmost importance is that the aim of any legislation should be the protection of the ecosystems integrity. This implies, first, the knowledge of the multitude of processes involved at different levels of organization, from cellular (bio-chemical), to whole-organism physiological responses, to population dynamics up to interactions within community and between it and the abiotic environment. The progress of our understanding of the ecosystem properties should then be translated in the development of operative tools to be integrated in the effects assessment procedure for an effective environmental protection and management.

Submitted on invitation.
Accepted on 13 February 2002.

REFERENCES

1. European Commission. *Technical guidance document in support of Commission Directive 93/67/EEC on risk assessment for new notified substances and the Commission Regulation (EC) 1488/94 on risk assessment of existing substances*. Part III. Brussels: European Commission; 1996.

2. Scientific Advisory Committee on Toxicity and Ecotoxicity of Chemicals of the European Commission-Ecotoxicity Section. *EEC water quality objectives for chemicals dangerous to aquatic environments (List I)*. CSTE/EEC; 1994. (Reviews of Environmental Contamination and Toxicology, 137).

3. Hermens JLM. Quantitative structure-activity relationships of environmental pollutants. In: Hutzinger O (Ed.). *Handbook of environmental chemistry.* vol. 2E. Berlin: Springer Verlag. 1989; p. 111-62.

4. Verhaar HJM, van Leeuwen CJ, Hermens JLM. Classifying environmental pollutants. 1. Structure activity relationships for prediction of aquatic toxicity. *Chemosphere* 1992;25:471-91.

5. Organisation for Economic Co-operation and Development. *Guidance document for aquatic effects assessment*. Paris: OECD; 1995. (OECD environment monographs no. 92).

6. Van Leeuwen CJ, van der Zandt PTJ, Aldenberg T, Verhaar HJM, Hermens JLM. Application of QSARs, extrapolation and equilibrium partitioning in aquatic effects assessment. I. Narcotic industrial pollutants. *Environ Toxicol Chem* 1992;11:267-82.

7. Nendza M. *Structure-activity relationships in environmental sciences*. London: Chapman & Hall; 1998.

8. Hermens JLM. Electrophiles and acute toxicity to fish. *Environ Health Persp* 1990;87:219-25.

9. Veith GD, Broderius SJ. Rules for distinguishing toxicants that cause type I and type II narcosis syndromes. *Environ Health Persp* 1990;87:207-11.

10. Russom CL, Bradbury SP, Broderius SJ, Hammermeister DE, Drummond RA. Predicting modes of toxic action from chemical structure: acute toxicity in the fathead minnow (*Pimephales promelas*). *Environ Toxicol Chem* 1997;16:948-67.

11. Marchini S, Passerini L, Hoglund MD, Pino A, Nendza M. The toxicity of aryl- and benzylhalides to *Daphnia magna* and classification of their mode of action based on QSARs. *Environ Toxicol Chem* 1999;18(12):2759-66.

12. Lieth H. Ecosystem principles for ecotoxicological analyses. In: Schuurmann G, Markert B (Ed.). *Ecotoxicology*. New York and Heidelberg: John Wiley & Sons Inc. and Spektrum Akademischer Verlag; 1998.

13. Stephan CE. Increasing the usefulness of acute toxicity tests. Aquatic toxicology and hazard assessment. In: Pearson JG, Foster RB, Boishop WE (Ed.). *5th conference, ASTM STP 766*. American Society for Testing and Materials; 1982. p. 69-81.

14. Verhaar, HJM, De Wolf W , Dyer S, Legierse KCHM, Seinen W, Hermens JLM. An LC50 *vs* time model for the aquatic toxicity of reactive and receptor-mediated compounds. Consequences for bioconcentration kinetics and risk assessment. *Environ Sci Technol* 1999;33:758-63.

15. McCarty LS, Mackay D. Enhancing ecotoxicological modeling and assessment. *Environ Sci Technol* 1993;27:1719-28.

16. Kooijman SALM. Parametric analyses of mortality rates in bioassays. *Water Res* 1981;15:107-19.

17. Niesink RJM, de Vries J, Hollinger MA. *Toxicology: principles and applications*. Boca Raton, FL: CRC Press; 1996.

18. Pack S. *A review of statistical data analysis and experimental design in OECD aquatic toxicity test guidelines*. Sittingbourne, UK: Shell Research Ltd.; 1992.

19. Organisation for Economic Co-operation and Development. *Report of the final ring test of the* Daphnia magna *reproduction test*. Paris: OECD; 1997. (Environmental Health and Safety Publications Series on Testing and Assessment, 6).

20. Mallett MJ, Grandy NJ, Lacey RF. Interlaboratory comparison of a method to evaluate the effects of chemicals on fish growth. *Environ Toxicol Chem* 1997;16:528-33.

21. Marchini S, Tosato ML, Norberg-King TJ, Hammermeister DE, Hoglund MD. Lethal and sublethal toxicity of benzene derivatives to the fathead minnow, using a short-term test. *Environ Toxicol Chem* 1992;11:187-95.

22. Stephan CE, Rogers JW. Advantages of using regression to calculate results of chronic toxicity tests. In: Bahner RC, Hansen DJ (Ed.). *Aquatic toxicology and hazards assessment*. 8. Symposium ASTM STP 891. Philadelphia, USA: American Society for Testing and Materials; 1985.

23. Organisation for Economic Co-operation and Development. *Report of the OECD workshop on statistical analysis of aquatic toxicity data*. Paris: OECD; 1998. (Environmental Health and Safety Publications Series on Testing and Assessment, 107).

24. Kooijman, SALM. Process-oriented descriptions of toxic effects. In: Schuurmann G, Markert B (Ed.). *Ecotoxicology*. New York and Heidelberg: John Wiley & Sons Inc. and Spektrum Akademischer Verlag; 1998.

25. Slooff W, Canton JH, Hermens JLM. Comparison of the susceptibility of 22 freshwater species to 15 chemical compounds. I. (Sub)acute toxicity tests. *Aquat Toxicol* 1983;4:113-28.

26. Organisation for Economic Co-operation and Development. *Report of the OECD workshop on the extrapolation of laboratory aquatic toxicity data to the real environment*. Paris: OECD; 1992. (OECD Environment Monographs, 59).

27. Zeeman MG. Ecotoxicity testing and estimation methods developed under section 5 of the toxic substances control act (TSCA) In: Rand GM (Ed.). *Fundamentals of aquatic toxicology.* 2nd ed. Washington, DC: Taylor and Francis; 1995.

28. Blank H. Species dependent variation among aquatic organisms in their sensitivity to chemicals. *Ecological Bull* 1984;36:107-19.

29. Thurston RV, Gilfoil TA, Meyn EL, Zajdel RK, Aoki TI, Veith GD. Comparative toxicity of ten organic chemicals to ten common aquatic species. *Water Res* 1985;19:1145-55.

30. Sloof W, Van Oers JAM, De Zwart D. Margin of uncertainty in ecotoxicological hazard assessment. *Environ Toxicol Chem* 1986; 5:841-52.

31. Hoekstra JA, Vaal MA, Noteboom J, Slooff W. Variation in the sensitivity of aquatic species to toxicants. *Bull Environ Contam Toxicol* 1994;5:98-105.

32. Nendza M, Wenzel A. Statistical approach to chemical classification. *Sci Total Environ Suppl* 1993;1459-70.

33. Vittozzi L, De Angelis G. A critical review of comparative acute toxicity data on freshwater fish. *Aquatic Toxicol* 1991;19:167-204.

34. Sprague JB. Factors that modify toxicity. In: Rand GM, Petrocelli SE (Ed.). *Fundamentals of aquatic toxicology. Methods and applications*. Washington, DC: Hemisphere Pub. Corp; 1985.

35. Marchini S, Hoglund MD, Broderius SJ, Tosato ML. Comparison of the susceptibility of daphnids and fish to benzene derivatives. *Sci Total Environ Suppl* 1993;1:799-808.

36. Stephan CE, Mount DI, Gentile JH, Chapman GA, Brungs WA. *Guidelines for deriving numeric National Water Quality Criteria for the protection of aquatic organisms and their uses, PB-85-227049*. Duluth, Minnesota, USA: US Environmental Protection Agency; 1985.

37. Kooijman SALM. A safety factor for LC50 values allowing for differences in sensitivity among species. *Water Res* 1987;21:269-76.

38. Wagner C, Løkke H. Estimation of ecotoxicological protection levels from NOEC toxicity data. *Water Res* 1991;25:1237-42.

39. Aldenberg T, Slob W. Confidence limits for hazardous concentrations based on logistically distributed NOEC toxicity data. *Ecotoxicol Environ Safety* 1993;25:48-63.

40. Van Straalen NM, Denneman CAJ. Ecotoxicological evaluation of soil quality criteria. *Ecotoxicol Environ Safety* 1989;18:241-51.

41. Hopkin SP. Ecological implications of “95% protection levels” for metals in soil. *Oikos* 1993;66:137-41.

42. Pratt JR, Cairns J Jr. Ecotoxicology and the redundancy problem: understanding effects on community structure and function. In: Newman MC, Jagoe CH (Ed.). *Ecotoxicology: a hierarchical treatment*. Boca Raton, FL: CRC; 1996. p. 347-70.

43. Nendza M, Volmer J, Klein W. Risk assessment based QSAR estunates. In: Karcher W, Devillers J (Ed.). *Practical applications of quantitative structure-activity relationships (QSAR) in environmental chemistry and toxicology.* Dordrecht: Kluwer; 1990.

44. Okkerman PC, Plassche EJ, Slooff W, Van Leeuwen CJ, Canton JH. Ecotoxicological effects assessment: a comparison of several extrapolation procedures. *Ecotox Environ Safety* 1991;21:182-93.

45. Newman MC, Ownby DR, Mézin LCA, Powell DC, Christensen TRL, Lerberg SB, Andersen BA. Applying species-sensitivity distributions in ecological risk assessment: assumptions of distribution type and sufficient numbers of species. *Environ Toxicol Chem* 2000;19(2):508-15.