Journal of Hazardous Materials, 35 (1993) 255–271
Elsevier Science Publishers B.V., Amsterdam

Using toxicity data to evaluate ecological effects at Superfund sites*

P. Bruce Duncan

U.S. Environmental Protection Agency, Region 10, 1200 6th Ave, ES-098, Seattle, WA 98101-1128 (USA)

Abstract

Toxicity data (used here to mean toxicity tests or bioassays) have been employed in two ways at Superfund sites. In a retrospective manner, bioassays have served as direct measurement endpoints (e.g., bioassays of contaminated sediment) and, in a predictive manner, bioassay data from the literature have been used to interpret other measurement endpoints such as contaminant concentrations. Ecological assessments for several Superfund sites provide examples of how these two approaches have been used to understand the ecological effects of soil, surface water, and sediment contamination. It appears toxicity data can be particularly useful in several ways, from screening lists of contaminant concentrations for potential toxicity to evaluating the geographical extent of demonstrably toxic contamination.

1. Introduction

This symposium investigates the role of ecological assessments in managing chemical pollution from a variety of perspectives, from the regulatory mandates to the details of how to conduct one. Figure 1 shows the basic components of a general paradigm for conducting an ecological assessment [1] and a view of how the diverse array of topics being discussed at this symposium relates to these components.

When a chemical contaminant has been released, the goal of the ecological assessment is to provide risk managers with as scientific an evaluation of the effects as possible. The risk managers then weave this information together with other components such as human welfare, cost, benefit, etc. to decide on a course of action. Superfund site investigations are an area where ecological

* Paper presented at the Symposium on the Role of Environmental (Ecological) Assessment in the Management of Chemical Pollution, American Chemical Society 204th National Meeting, Division of Industrial & Engineering Chemistry, Washington, DC, August 26–28, 1992.

0304-3894/93/$06.00

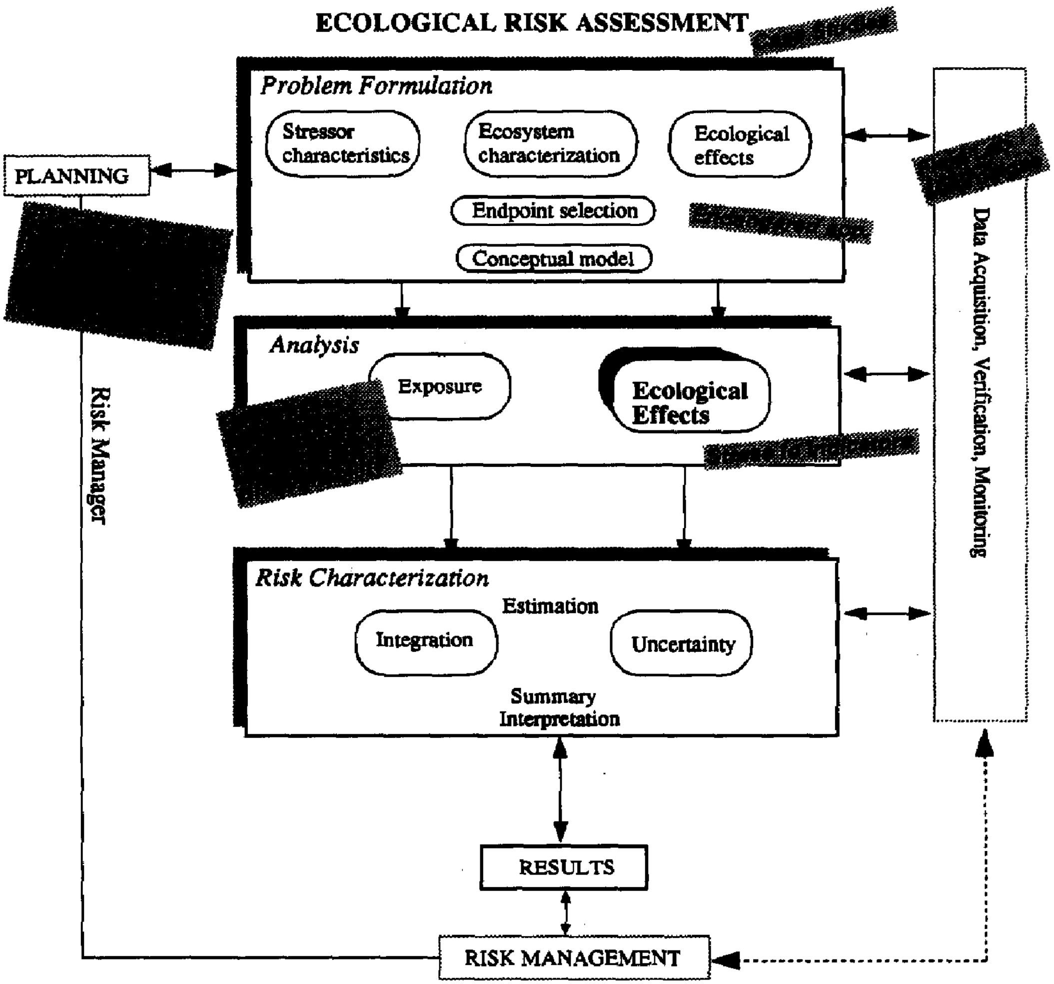

Fig. 1. General paradigm for conducting ecological risk assessments (taken from the framework proposed by the EPA Risk Assessment Forum [1]). Superimposed, in shaded boxes, are the array of topics of this symposium on the role of ecological assessments in managing chemical pollution. In this paper, the use of toxicity tests in the Analysis phase is discussed, where ecological effects are estimated.

assessments of chemical releases are becoming more commonplace. A variety of approaches to completing assessments has been attempted at these sites and the potential now exists to evaluate this burgeoning data base and judge which assessments have been successful. A successful assessment not only evaluates the threat to ecosystems from releases, but also provides the basis and justification for a course of action and clean-up levels when possible.

In the paradigm shown in Fig. 1, the main components of the ecological assessment are problem formulation, exposure and effects analysis, and risk characterization. Here, only a small part of this paradigm is examined, specifically, the use of toxicity tests to evaluate ecological effects in the analysis step.

Toxicity tests (measuring the response of organisms exposed to contaminants relative to a control) are important in characterizing effects, particularly the relationships between stressors and the ecological responses they may elicit [1]. Because of this importance, information from these tests can be critical in completing an overall risk assessment. This paper discusses how toxicity tests have been used in two ways at Superfund sites. Retrospectively, they have served as direct measurement endpoints (e.g., testing contaminated sediment using amphipod crustaceans) and, predictively, published data have been used to evaluate other measurement endpoints such as contaminant concentrations (often termed the quotient method).

1.1 Endpoints

A very important step in problem formulation is the selection of endpoints for assessment and measurement. Descriptions of these two types of endpoints are given by Suter [2]. Briefly, assessment endpoints are what we care about (for example, a decline in a local run of wild salmon) and measurement endpoints are the actual data we collect (e.g., testing the toxicity of the water in the salmon stream by using a surrogate such as hatchery salmon) that we then relate to the assessment endpoint. In some cases, the measurement and assessment endpoints can be the same. For the salmon example above, one could assess a decline by measuring the population directly. In other cases, the two types of endpoints will differ. For example, a model could be used to predict uptake of a contaminant into raptor eggs based on body burdens of the raptor's prey (measurement endpoint) with the aim to evaluate the effect on nesting success (assessment endpoint). Although the choice of endpoints appears to be limitless, the types of measurements that are likely to be made can be placed into a relatively few broad categories such as concentrations of contaminants and their transformation products in various media (including tissues), bioassays (*in situ*, laboratory, etc.), population or community level studies, and biomarkers (e.g., proteins manufactured in response to stress).

1.2 Ecotoxicity data

Ecotoxicology is the science that "seeks to predict the impacts of chemicals upon ecosystems" [3]. It is a relatively new and developing field, which, in its simplest form, is the application of toxicological principles to ecosystems (associations of biotic communities and their abiotic habitats including structural components and functional interactions). In this approach, the nature and effects of pollutants are evaluated in terms of their sources, uses and properties, their physical and chemical interactions with the environment, and their effects on organisms and ultimately on ecosystems [4]. The Organization for Economic Cooperation and Development (OECD) has suggested a set of minimum data needed to assess effects of chemicals in the environment [5]. One of the data categories needed is ecotoxicity data. OECD lists fish mortality, impaired reproduction in crustaceans, and algal growth inhibition as examples of this type of data. Similarly, the U.S. Environmental Protection Agency

TABLE 1

Examples of the effect endpoints in the AQUIRE data base

Category	Effect
Cellular	Change in organelle structure Cytogenetic effect: Changes in the RNA and DNA
Tissue	Presence of physical damage to tissues
Developmental	Change in the ability to grow to a more mature life stage Change in timing between separate life stages such as metamorphosis, molting, emergence, yolk absorption
Behavioral	Avoidance or attraction to a chemical gradient
Bioconcentration	Accumulation of a toxicant in the tissues of the test organism compared to the water concentration
Individual	EC_{xx}: concentration that affects *xx*% of the tested organisms. Effects include abnormalities, detachment, development, enzyme activity, growth, pigment change, population size, pupation, reproduction, shell valve closure, uptake LC_{50}: concentration lethal to 50% of the tested organisms
Population	Abundance: number of organisms within the same species changes Change in cell number of algal species
Community	Change in number of species groups in a given community (species diversity)

(EPA) requires collection of toxicity data for a variety of taxonomic groups to derive water quality criteria to protect aquatic organisms [6]. In general, at least two major types of toxicity tests are conducted. A single contaminant can be tested over a range of concentrations such as a dilution series of cadmium chloride. This approach is useful in developing criteria. Alternatively, field samples of water, soil, or sediment containing suites of contaminants, perhaps unknown, can be collected and tested. In these tests, comparisons are made with some reference sample (presumed to be "clean" or to represent upgradient or local background conditions) or an array of samples collected along a suspected gradient of contamination.

Standardized bioassay results are often used to extrapolate from laboratory effects to estimate effects in the environment. Although single-species and single-contaminant toxicity tests cannot reliably predict effects of contaminants on an ecosystem, they can be integrated with other approaches such as microcosm tests, modelling, or field studies [3, 7]. Despite the uncertainty involved in extrapolating from toxicity tests to environmental effects [8], ecotoxicity data can be used effectively in ecological assessments.

Many types of bioassays can be conducted and numerous summaries describing available tests exist (see Dinnel's bibliography on sediment toxicity testing [9], for example). Tests can be conducted in the laboratory or the field, involve single-species, communities, or microcosms, and have endpoints from enzymatic changes to lethality or population growth effects. A data base of aquatic bioassay data that amply illustrates this wide range in endpoints is the AQUIRE (AQUatic Information REtrieval) data base maintained by the EPA, Duluth laboratory [10]. Table 1 lists some examples of the endpoints in the AQUIRE data base.

2. Direct assessment using toxicity tests

Toxicity tests can be used to determine the areal extent of contamination in a biologically meaningful way, to verify suspected toxicity, and to overcome uncertainties in relating concentration data to literature-derived values. Relatively simple ecological assessments may not include toxicity tests but instead rely on the evaluation of concentration data (comparison with literature values and criteria) coupled with site visit information [11, 12]. When toxicity tests are conducted early in the ecological assessment, they can be used to fill in data gaps and address uncertainties in evaluating the concentration data set. A particularly good example of the use of toxicity tests is the "triad" approach developed for marine sediments [13, 14]. In this approach, sediment chemistry, sediment bioassays, and changes in the benthic community are evaluated. These values can be scaled using reference values so that comparisons can be made among a variety of stations [15].

Several related problems can occur when toxicity tests are conducted on field samples. The first is how to select a reference area that ideally matches the site. For example, marine sediments should be matched for grain size, organic content, depth, hydrologic regime, etc. This may not be simple if the site includes a range of sediment types such as inter- and subtidal sands and muds. Because of such heterogeneity, several reference sites may be needed. Sometimes, no reference site is available so gradients are evaluated by comparing the site samples with "upstream" and "downstream" samples. Another consideration is how to assess the contribution from a site release given the presence of other sources of contamination. In this case, one will have to decide whether a truly uncontaminated background is a suitable reference and perhaps make a trade-off between understanding the overall existing risk versus the incremental risk due to the site release. An additional problem occurs if mortality in the controls (utilizing "clean" sediment, water, or soil) is excessive.

Choosing which species to use for the test may not be simple. A species that is too sensitive could die in all the samples collected in the gradient approach mentioned above. Conversely, an insensitive species may not help one distinguish moderately toxic areas from nontoxic areas. Choosing a species that is appropriately sensitive and somewhat representative of local species, will be a much easier process if the assessment endpoints have been identified.

Given the problems discussed above, before any toxicity tests are run, it is imperative to meet with the risk managers to determine how the test results will be used to make decisions regarding further testing, remediation, etc. Part of this determination includes defining what is an ecologically significant effect. For example, which of the following will be considered unacceptable: Ten percent mortality (how about 50%?); a statistically significant depression in diversity, growth, reproduction, etc.; or, a statistically significant effect for two out of three tests?

2.1 Delineate the extent of toxicity

Toxicity tests can be used to help delineate the extent of suspected toxicity in a geographical area. At some Superfund sites, an effective approach has been to evaluate chemical concentrations in an area, determine the gradients of contamination, then use toxicity test results from samples taken along the gradient to determine the extent of contamination based on biological effect. This method is useful when the contaminants are few, when it is cost-effective to determine the pattern of distribution on site (for example, if metals can be rapidly analyzed on site using X-ray fluorescence [16]), and when gradients are readily discernible. For example, in a wetland area contaminated by heavy metals, this approach was used to determine the gradients of contamination and then to select sampling sites for bioassays. Based on the measured relationship between concentrations and toxicity, the area of ecological concern (i.e., contaminated sediments capable of eliciting toxic effects) was delineated [17].

2.2 Verify suspected or predicted toxicity

A preliminary assessment of contamination usually involves evaluating available environmental information. Most commonly, this will be a listing of contaminant concentrations found in various media. In some cases, it will be possible to compare these concentrations with some sort of reference concentration and predict toxicity. For example, concentrations in a ditch, stream, etc., could be compared with the EPA water quality criteria [18]. This comparison readily yields a list of contaminants suspected as capable of exerting toxic effects. A somewhat different approach can be used for marine sediments in the Puget Sound, WA, area, where threshold effects concentrations have been calculated for a variety of contaminants [19]. Toxicity at a particular site may also be suspected based on descriptions of visible environmental effects. In some cases, this may be very obvious such as fish or bird kills, or the loss of vegetation due to a smelter discharge (Bunker Hill, ID, for example [20]). In other cases, the visible damage may be slight, and there may be effects on the habitat that are not related to a chemical release (cattle use, for example [11]). In addition to using reference concentrations, an attempt can be made to use toxicity data from the literature to predict what the toxic effects of environmental concentrations might be. By compiling this information early in the assessment process, one can decide whether to conduct toxicity tests to verify suspected toxicity if measured concentrations are within some pre-selected factor of the comparison concentrations.

2.3 Overcome uncertainties in evaluating concentration data

The availability of criteria help the ecological assessment tremendously but there are constant shortcomings. In many instances, there is no criterion against which to compare a contaminant concentration. Although criteria for water are the best developed of all the media, of the 189 parameters listed in the 1991 summary of ambient water quality criteria [21], only about 109 have chronic or acute values for marine or freshwater ecosystems. A typical scan of a water sample (using the Superfund target compound list [22, 23]) could include measurements of over 125 organic and 24 inorganic contaminants. Clearly, even the best collection of criteria available to date will often be inadequate.

Efforts to evaluate the effects of contamination in sediments have improved lately with the use of Apparent Effects Thresholds (AETs) [19], equilibrium partitioning [24], or the summary of toxicity tests by Long and Morgan [25]. Despite this progress, many uncertainties and limitations remain when using these approaches to evaluate marine sediment concentrations (see Table 2 for a summary). Evaluating freshwater sediment concentrations remains an even more difficult task due to the lack of criteria and lack of accessible data bases. Perhaps the most difficult medium at present to evaluate using concentration

TABLE 2

Comparison of three approaches used to evaluate marine sediments. A brief description and some of the limitations are given[a]

Method	Description and limitations
AET [19]	Apparent Effects Threshold — the concentration above which biological effects are always seen for a data set of matched sediment chemistry and effects – Requires a lot of matched data and is computed for a specific region (e.g., Puget Sound, WA) – Can be influenced by correlated contaminants
EP [24]	Equilibrium Partitioning — predicts pore water concentrations for single chemicals which are then compared with water quality criteria – Assumes water quality criteria protect infauna – Restricted to non-polar organic contaminants, so far – Need to measure the sediment organic content
ER-L ER-M [25]	NOAA's Effects Range Low and Medium — based on the 10th and 50th percentiles of matched effects and concentration data from many sources, including the two described above. Data were screened based on the concordance between concentration and effects – ERLs and ERMs are influenced by the type and amount of data. The relative degree of confidence of these measures are given for each analyte – Freshwater and marine data are not evaluated separately

[a]Each of these shares the limitation that not all contaminants have been evaluated. A more detailed synopsis of these and other approaches is given in [25].

data is soil. Although the PHYTOTOX data base (a compilation of data concerning toxicity mainly of pesticides [26]) is somewhat accessible, it too shares many of the problems associated with using data bases (see Section 3.1, below).

Even when criteria are available, if several contaminants are just below the level of concern, their cumulative effects, which can be additive, synergistic, or antagonistic, may not be predictable. Another concern is whether criteria apply equally well at all sites. Indeed, the EPA regulations provide for the development of site-specific water quality criteria [27]. In a somewhat similar fashion, the State of Washington, Department of Ecology allows exceedances of marine sediment standards to be "challenged" with biological information before remedial action needs to be undertaken at contaminated sites. The results of biological testing (bioassays and/or benthic community analysis) can override the criterion exceedance [28]. In general, whenever toxicity is estimated, direct assessment may be needed because actual exposures, bioavailability, matrix effects, etc. have not been addressed.

3. Predicting toxicity using literature data

Despite the limitations discussed above, toxicity test data reported in the literature have been used successfully to evaluate contaminant concentrations. Here, two approaches are discussed that are used to examine data for water, sediment, and soil. One approach is to evaluate the medium as habitat. For example, one can ask whether the measured concentration of zinc in soil (e.g., μg Zn/kg of soil) could have an adverse effect on earthworms that live in and ingest the soil particles. Similarly, one may need to know if phenanthrene in marine sediments could affect benthic crustaceans that burrow in the sediments. Or, one could question whether a certain concentration of dichloroethylene (DCE) in a lake might affect the ability of fish to reproduce. A second way to estimate effects is to treat the contaminant in the water or soil as a dose (e.g., mg of contaminant/kg body weight/day). In this approach, one might evaluate the potential effect of DCE in lake water on a mule deer drinking the water, the effect of zinc in the soil on a field mouse that ingests the soil while feeding and preening, or the effect of chlordane in freshwater sediments on waterfowl incidentally ingesting the sediment while foraging on rooted aquatic vegetation. Clearly, the approach selected is determined in large part by the organism to be evaluated.

3.1 Predicting toxicity of water, sediment, and soil as habitat

Some suggestions are given here on how to use literature values to evaluate the effect of contaminant concentrations on organisms inhabiting different media. In many cases, searchable data bases do not exist and the available literature will have to be scoured on a contaminant-specific basis.

The major shortcoming to comparing site concentrations with literature values or criteria is the lack of either or both of these for many chemicals.

Other problems using literature values include lack of relevance to the site (available data may be for a species such as the goldfish, while the species of concern on site are salmonids), having information on only one or a few species (one would like many tests on a variety of species from different major taxonomic groups, similar to the approach for deriving water quality criteria), and the lack of information on mixtures of contaminants. Nevertheless, the data may be particularly valuable for early screening of contaminants during the ecological assessment. One relatively easy approach is to compare the maximum measured concentration against the smallest literature toxicity value for the most sensitive species and include modifying factors to account for the uncertainty in extrapolating across species and endpoints. Although this approach may not reflect reasonable conditions, confidence that a contaminant is *not* toxic may be high if the maximum site concentration falls below the modified literature value. In some cases it may be appropriate to back off from this conservative first cut. For example, when salmonid fish have been selected as assessment endpoints, it may be appropriate to use only data for anadromous fish. Other considerations may lead to using average or 95th percentile site concentrations instead of the maximum. For instance, when chronic toxicity is being evaluated for a species that ranges across the site, average concentrations may be appropriate. This type of decision should only be made once the assessment endpoints have been defined (see Section 1.1).

In no case should the conservatism in the screening approach suggested above be overadjusted to the point that one assumes reality is being modeled. There are simply too many unsupported assumptions inherent in applying lab studies on single, surrogate species, responding to a single contaminant, to the more complex, site-specific, multiple-contaminant, field situation. Concentrations that exceed the ultra-conservative screening levels do not necessarily indicate environmental harm. For exceedances in, say, the range of one order of magnitude, adjustments to the approach that remain conservative can be useful in further screening the list of contaminants of concern.

3.1.1 Water

As discussed previously (Section 2.2), the easiest way to evaluate water concentrations is by comparing them to the water quality criteria (which are derived by distilling large amounts of toxicity test data [6]). The water quality criteria may not be protective of benthic species, so sediment toxicity may require a separate evaluation. Similarly, the criteria may not account for toxicity to wildlife such as waterfowl, and other approaches must be used for these species. In the absence of criteria, the AQUIRE data base [10] can be useful in retrieving toxicity and bioaccumulation information on specific chemicals.

3.1.2 Marine and freshwater sediment

In addition to the approaches described below, the primary and "gray" literature can be consulted on a site-specific basis for information on sediment toxicity. Table 2 describes three of the approaches to predicting the toxicity of marine sediment based on concentration data and outlines some of the limitations of each approach. No searchable data base similar to AQUIRE exists yet for marine sediments; the most comparable effort is the compilation by Long and Morgan [25]. Freshwater sediment concentrations can be evaluated for some contaminants by using several comparison numbers. Some attempts have been made to collate Ontario guidelines, Great Lakes classification, Long and Morgan's compilation, and other sources of "reference" numbers (see, for example [29]).

3.1.3 Soil

No criteria are available for soil as habitat. Information from the PHYTOTOX data base [26] can be used to evaluate phytotoxicity based on contaminant concentration. Some information can be gleaned from the US Fish and Wildlife Service documents on approximately 20 contaminants [30], and the EPA sludge documents [31].

3.2 Predicting toxicity of water, sediment, and soil using dose–response toxicity test data

This approach is very similar to that used in human health risk assessment. Dose–response data are coupled with exposure assumptions to evaluate the toxicity of a contaminant. First the species of concern must be selected (e.g., mule deer, red-tailed hawk, field mouse). Then, many assumptions must be made concerning the exposure of the organism to the contaminant. For example, some of the assumptions for a mule deer ingesting soil include: the amount of soil ingested through each likely pathway (e.g., soil on vegetation), the amount of the contaminant that is bioavailable, the mule deer's body weight, and the amount of the deer's total soil ingestion that is from contaminated soil (also called an area use factor). The exposure calculations result in an estimate of dose in units such as mg-contaminant/kg-body weight/day. The next step is to compare this dose with available literature data. If data specifically for mule deer are not available, values for other ungulates including domesticated species can be evaluated. There is no formal protocol yet for making this comparison in an ecological assessment, but the same considerations occur as in a human health assessment (e.g., Should an uncertainty factor be applied to reflect the use of cow data for mule deer? How should different endpoints be evaluated when they range from decreased growth to a change in blood chemistry to mortality?). Similarly, the following issues should also be addressed: Have multiple pathways been included and how were they integrated? Is carcinogenicity considered an endpoint of concern, or is it assumed that in wild populations organisms die before cancers become debilitating? Does the assessment account for sensitive life stages or other stresses not related to contaminant releases that may increase susceptibility to the contaminant?

Despite these problems in predicting doses and extrapolating effects from non-site-related studies, conservative assumptions can be made and this approach can be successfully used to screen contaminants and help decide whether further investigation is warranted. Before scouring the literature for the most realistic exposure data, one can (1) use exaggerated intake values; (2) assume all the contaminant is bioavailable; (3) assume the organism ingests the contaminant at the maximum concentration every day; (4) use the literature toxicity value for the most sensitive species; and, (5) apply uncertainty values for each extrapolation (species to species, lab to field, etc.). Although this approach will not reflect reality, confidence that a contaminant is *not* toxic may be high if the calculated dose falls below the toxic literature dose.

There are at least two problems that can occur with this approach. First, excessive zeal in using ultra-conservative values may result in no contaminants being screened out of the risk assessment. So, as suggested in Section 3.1, above, it may be appropriate to refine the calculations as better exposure or toxicity data are developed. This will reduce the magnitude of uncertainty, while the conservatism is controlling the direction of uncertainty. Second, when the calculated dose exceeds the literature dose, a prediction that the measured concentrations are toxic is likely to be incorrect especially when the two doses are within an order of magnitude of each other (because of the conservatism). This suggests, therefore, that this approach be used, not to model reality, but to screen contaminants using moderately to extremely conservative assumptions. Field studies can then be used to evaluate any suspected exposure or toxicity. As this approach continues to be used, some generalities may emerge. For instance, at an abandoned mine site where water was contaminated by metals, the toxicity was predicted to be greater to freshwater aquatic organisms living in the water than to terrestrial biota that might drink the water [11].

4. Comparison of use at several Superfund sites

Provided below are some examples of how bioassays and literature toxicity data have been used to evaluate contaminated media at several Superfund sites in EPA's Region 10. These sites, located in the states of Washington and Idaho, were chosen to illustrate the actual use of bioassay data to evaluate contaminated water, sediment, and soil. These sketches are presented to give an idea of the assumptions, specific methods, and decisions that were reached, and are not intended to represent ideal or detailed case studies. As more ecological risk assessments are completed for Superfund sites, successes and failures of the approaches mentioned here can be better scrutinized.

4.1 Bunker Hill, Idaho — Soil, water, and sediment contamination (abstracted from [20])

At the Bunker Hill site, soils became contaminated by fallout from emissions from a lead smelter. Phytotoxicity of soils was evaluated by comparing metals concentrations in soil with literature-derived values. This comparison identified several metals as contaminants of concern. Toxicity to soil biota was similarly evaluated. Toxicity to mammals and birds was evaluated by calculating assumed doses and comparing these with available dose–response literature. Some of the assumptions used included: lifetime exposure for deer and waterfowl (versus ranging off-site); total absorption of contaminants; assumptions on how much vegetation, soil, and water are ingested by a mouse, mallard, and deer; and, a 1% concentration factor from soil into vegetation. Based on soil and water concentrations, a range of intakes was calculated (the range was over a factor of about four). In utilizing toxicity data, lethal and sublethal endpoints were evaluated. Lowest observed adverse effect levels or lowest observed lethal doses were selected, when possible, as the reference lethal toxicity values, but most values were LD_{50}s (lethal to half the tested organisms at that dose). For contaminants with more than one valence state, toxicity data for the most toxic state were selected. Toxicity data for deer in particular were lacking (only arsenic and mercury were evaluated), but sufficient information was available to evaluate nine metals (but not mercury) for toxicity to mice and six metals for toxicity to waterfowl.

Mine tailings on-site have affected the South Fork of the Coeur d'Alene River. Water seeping through tailings and the deposition of tailings into the river have been the main exposure pathways. In the 1930's, assay tests indicated the river water was lethal to native fish species and surveys found stretches of the river with no life. More recent surveys have found small populations of fish. Over the years, a variety of toxicity tests have been conducted. Between 1973 and 1988, river water was tested using the seven-day *Ceriodaphnia* life cycle, Rainbow trout *in situ*, cutthroat trout, fathead minnow, and algal growth tests. These showed different sensitivities among the organisms (the *in situ* test using the rainbow trout was most sensitive — 100% mortality within 96 hours at all stations within the site) but all indicated significant toxicity from either seeps reaching the river or the river water itself. Water quality criteria were used as part of the screening criteria to identify the contaminants of concern and, later, to identify cadmium, lead, and zinc as the contaminants in the water with the most detrimental concentrations (exceedances by factors of approximately 15 to 17 times the criterion). Degraded water quality was evident both from this analysis and the freshwater bioassay results. Water as drinking water was combined with the soil ingestion pathway for evaluation of effects on wildlife.

Freshwater sediments were tested with the crustaceans, *Daphnia and Hyalella*. Results were not conclusive (*Daphnia* numbers were reduced, but *Hyalella* increased) and no chronic test was conducted. Despite these poor assay data, effects due to the seep discharges and tailings were evident in studies showing reduced diversity of benthic organisms in the river. Freshwater sediment concentrations were compared with sediment concentrations from other nearby rivers rather than with literature-derived or other benchmark values.

4.2 Eagle Harbor, Washington — Marine sediment contamination (abstracted from [32])

Marine sediments in Eagle Harbor were contaminated by mercury (from shipyard operations) and polynuclear aromatic hydrocarbons (PAHs; from a creosote treatment facility). Chemical analyses of the sediment were used to describe the extent of contamination. Bioassays were conducted on sediments collected from a subset of the chemistry stations and included the oyster larvae and the haustorid amphipod acute assays. Some previous bioassay work with the amphipods had revealed a particularly "hot" spot. Although there were some problems with the controls, results were adjusted to account for mortality of the oyster larvae and variations in sediment grain-size distribution among the samples. Sediment concentrations were compared with Apparent Effects Thresholds to define areas of concern, and smaller, final problem areas were selected based on the bioassay results. Because only a small number of samples were collected across what turned out to be a wide variety of habitats, the data concerning effects in benthic communities were somewhat inconclusive, but showed no gross changes in major taxonomic groups other than enhancement of polychaetes at some stations and the effect of the hot spot.

4.3 McChord Air Force Base, Washington — Freshwater and sediment contamination (abstracted from [12])

The major contaminants of concern were organics (trichloroethylene and dichloroethylene) in ground water. Habitats selected for ecological evaluation were the ponds and wetlands potentially impacted by the ground-water plume. On the basis of ground-water contamination, surface water and sediments were sampled from five lakes on site.

Freshwater in the ponds was evaluated using water quality criteria. Acute and chronic criteria were used. For cadmium, mercury, and some pesticides, the criteria were below the analytical detection limit. For some contaminants, toxicity data were abstracted from the specific water quality criteria documents. These data were used to prepare curves representing the cumulative number of species affected as contaminant concentration increased. These curves were used to evaluate exceedances of criteria in terms of the expected percent of aquatic species in each lake likely to be affected. When chronic values were lacking they were estimated from acute values using a factor of 20. For four organic compounds there were no criteria, so lowest-observed-effect concentrations were obtained from AQUIRE and used to represent chronic threshold concentrations.

Potential risk from drinking water and ingesting lake sediment by terrestrial organisms and ducks was evaluated using the dose method. Assumptions included: total measured metals were all bioavailable; depuration was zero; and, ingestion occurred at the most contaminated site. When necessary, chronic doses were estimated from acute doses by dividing lethal doses by 1000. Acceptable daily intakes were taken from lowest observed effect levels divided by 10. Incidental sediment ingestion rates were estimated as one percent of total ingestion except for ducks where 10 percent was assumed. Data were lacking for zinc.

The equilibrium partitioning method was used to estimate the toxicity of nonpolar organics in sediments to the interstitial water. To do this, a percent organic carbon was assumed for the sediments in order to normalize the chemical concentrations. Predicted risks to aquatic life varied by chemical, medium, and lake. Results indicated greater toxicity predicted in interstitial water than in surface water. No toxicity was predicted to terrestrial animals or birds on site.

5. Other uses of ecotoxicity literature data

5.1 Risk-based detection limits

Considerable effort has been made to improve the quality of data generated as part of Superfund risk assessments (see, for example, EPA's recent guidance on data usability [33]). Analytical detection limits have improved greatly, so it is now practical to estimate the level of detection needed in an environmental sample that will allow the risk-based screening of the measured concentrations. Toxicity data can be used to help define these risk-based detection limits. Perhaps the easiest approach is to use existing criteria or comparison numbers and ask the laboratory for analyses that can achieve these concentrations. One could also use dose–response data and some exposure assumptions to estimate concentrations of concern in various media. These concentrations could then be compared with laboratory capabilities. This approach would provide data the risk assessor can use and would bring to discussion, at a very early stage, any problems with the requested detection limits (for example, a request for measurement of dioxin at or below the water quality criterion may exceed laboratory analytical ability).

5.2 Determine the need for toxicity testing, type of tests, and likelihood of success

Evaluations based on literature data may be inadequate or unsatisfactory (see the previous example of lack of toxicity data for deer in Section 4.1 and information on zinc in Section 4.3, above). Nevertheless, for some contaminants, although conclusions cannot be drawn, sufficient information may be available to determine whether bioassays should be conducted, which ones are likely to be appropriate, and what is the likelihood of success. For example, one could decide that bioassays will be conducted if ambient concentrations are greater than one tenth the literature-derived concentration or if the data base for the literature-derived value does not include taxa of concern (or surrogates) identified in the endpoint selection process. If a variety of taxa is represented, ranking them by sensitivity to the contaminants could help one choose from the available assortment of tests. Species that appear to have low variability in their response to a contaminant can be selected preferentially.

The development of a data base containing the results of bioassay tests that have been conducted at Superfund sites (see, for example, efforts by NOAA in this direction [34]) will prove to be invaluable to ecologists and managers having to choose whether to conduct tests and which ones to conduct.

6. Conclusions

Toxicity data play a major role in ecological risk assessment. Certainly, criteria are useful for screening or even conclusive "desk-top" assessments. Conducting bioassays *in situ* or on collected samples of soil, sediment, and water can provide strong evidence of effects of contaminants. Examples from several Superfund sites indicate the types of problems involved with using these data. Even though the protocols for the tests themselves are standardized, there was no standardization concerning the number and types of tests conducted for these sites, or standardization of how the results were used to make decisions. Several useful approaches discussed here (MTCA [28] and the water quality criteria development guidelines [6]) may help provide more consistency. Also in a relatively early state of development is the ability to predict effects from environmental concentrations by using literature-based approaches. The Superfund sites discussed show the wide ranges in assumptions used and provide an example of the ability to use conservative values and reach conclusions. Some suggestions are presented for predicting the toxicity of contaminated soil, sediment, and water either as habitat or as an ingested dose. It is therefore, recommended to expand the effort begun by NOAA to compile the existing bioassay information from Superfund sites into an accessible, usable, data base. This will be a good first step in promoting consistency, improving the success of toxicity testing, and providing better information to risk managers

Acknowledgments

Many of the ideas presented here have benefitted greatly by discussions with members of the EPA Region 10 and other regional EPA Biological Technical Assistance Groups, particularly Dr. Pat Cirone and Doug Beltman. D. Beltman, P. Cirone, and Anne Sergeant provided critical reviews of the manuscript. Disclaimer: This article does not represent the official views of the U.S. Environmental Protection Agency.

References

1 U.S. Environmental Protection Agency, Framework for Ecological Risk Assessment, Risk Assessment Forum, EPA/630/R-92/001. U.S. EPA, Washington, DC, 1992.

2 G.W. Suter II, Endpoints for regional ecological risk assessments, Environ. Manage., 14 (1) (1989) 19–23.
3 S.A. Levin, M.A. Harwell, J.R. Kelly and K.D. Kimball, Ecotoxicology: Problems and Approaches, Springer-Verlag, New York, 1989, 547 pp.
4 D.W. Connell and G.J. Miller, Chemistry and Ecotoxicology of Pollution, Wiley, New York, 1984, 444 pp.
5 OECD, Guidelines for Testing of Chemicals. Organisation for Economic Cooperation and Development, Publications Office, Paris, France, 1981 .
6 U.S. EPA, Guidelines for deriving numerical national water quality criteria for the protection of aquatic organisms and their uses, Washington, DC, 1985.
7 J.M. Neuhold, Toward a meaningful interaction between ecology and aquatic toxicology, In: T.M. Poston and R. Purdy (Eds.), Aquatic Toxicology and Environmental Fate, Vol. 9 ASTM STP 921. ASTM, Philadelphia, PA, 1986, pp. 11–21.
8 M.A. Harwell and C.C. Harwell, Environmental decision making in the presence of uncertainty, In: Ecotoxicology: Problems and Approaches, S.A. Levin, M.A. Harwell, J.R. Kelly and K.D. Kimball (Eds.), Springer-Verlag, New York, 1989, pp. 517–540.
9 P.A. Dinnel, Annotated bibliography of bioassays related to sediment toxicity testing in Washington State. Final report from Fisheries Research Institute, School of Fisheries, University of Washington, Seattle, WA, to Seattle District, US Army Corps of Engineers, Seattle, WA, FRI-UW-9017, 1990, 126 pp.
10 A. Pilli, D.O. Carle and B.R. Sheedy, AQUIRE: Aquatic toxicity information retrieval data base. A technical support document, U.S. EPA Environ. Res. Lab., Duluth, MN, 1989, pp. 11–14.
11 U.S. EPA, Remedial investigation report: Silver Mountain Mine, Okanogan County, WA, U.S. EPA, Region 10, Seattle, WA, 1990.
12 U.S. Army Corps of Engineers, Ecological risk assessment of McChord Air Force Base — Area C/American Lakes Garden Tract, by Ebasco Environmental, Bellevue, WA, for Seattle District U.S. Army COE, 1991, pp. 1–1 to 5–3.
13 E.R. Long and P.M. Chapman, 1985, A sediment quality triad: Measures of sediment contamination, toxicity and infaunal community composition in Puget Sound, Mar. Pollut. Bull., 16 (1985) 405–415.
14 P.M. Chapman, R.N. Dexter and E.R. Long, Synoptic measures of sediment contamination, toxicity and infaunal community composition (the Sediment Quality Triad) in San Francisco Bay, Mar. Ecol. Prog. Ser., 37 (1987) 75–96.
15 E.R. Long, The use of the Sediment Quality Triad in classification of sediment contamination, In: Contaminated Marine Sediments — Assessment and Remediation, National Academy Press, Washington, DC, 1989, pp. 78–99.
16 R.G. Henry, D.W. Charters and L. Kaelin, The use of X-ray fluorescence to field screen sediment for metal concentrations prior to performance of a toxicity evaluation, In: Society of Environ. Toxicol. and Chem., 12th Annu. Meeting, Abstract 402, Washington, DC, 1991, p. 101.
17 D.W. Charters, R.G. Henry, E. Helmer and D. Beltman, The use of a chronic sediment toxicity evaluation to delineate the remedial boundary at a Superfund site, In: Society of Envir. Toxicol. and Chem., 12th Annu. Meeting, Abstract 545, Washington, DC, 1991, p. 136.
18 U.S. EPA, Quality criteria for water, EPA 440/5-86 001, Washington, DC, 1986.
19 U.S. EPA, Briefing report to the EPA science advisory board: The apparent effects threshold approach, by PTI Environmental Services, Bellevue, WA, for U.S. EPA, Region 10, Seattle, WA, 1988, 57 pp.
20 U.S. EPA, Ecological risk assessment for the Bunker Hill Superfund site, by SAIC, Bothell, WA, and EP & T, Corvallis, OR, for U.S. EPA, Region 10, Seattle, WA, 1991.
21 U.S. EPA, Water quality criteria summary (chart). Office of Sci. and Technol., Washington, DC, 1991.
22 U.S. EPA, Contract laboratory program statement of work for inorganic analysis: Multimedia, multi-concentration, Document No. ILM01.0 Office of Emergency and Remedial Response, Washington, DC, 1990.

23 U.S. EPA, Contract laboratory program statement of work for organic analysis: Multimedia, multi-concentration, Document No. OLM01.2. Office of Emergency and Remedial Response, Washington, DC, 1991.
24 U.S. EPA, Interim sediment quality criteria values for nonpolar hydrophobic organic contaminants. Office of Water Regulations and Standards, Washington, DC, 1988.
25 E.R. Long and L.G. Morgan, The potential for biological effects of sediment-sorbed contaminants tested in the national Status and Trends Program. NOAA, National Ocean Service, Seattle, WA, 1990.
26 C.L. Royce and J.S. Fletcher, PHYTOTOX user's manual, University of Oklahoma, Ada, OK, 1982.
27 40 CFR Parts 131.11 and 131.13, Washington, DC, 1992.
28 Washington State Dept. of Ecology, The Model Toxics Control Act Cleanup Regulation, Chapter 173-340 WAC, Olympia, WA, 1991, 133 pp.
29 J. Bennet and J. Cubbage, Summary of criteria and guidelines for contaminated freshwater sediments. Washington State Dept. of Ecology, Olympia, WA, 1991, 9 pp.
30 U.S. Fish and Wildlife Service, Contaminant Hazard Review Series, Biol. Reports 85(1.1) to 85(1.19), by R. Eisler, Patuxent Wildlife Research Center, Laurel, MD, 1985–1989.
31 U.S. EPA, Environmental profiles and hazard indices for constituents of municipal sludge. Office of Water Regulation and Standards, Washington, DC, 1985.
32 U.S. EPA, Final remedial investigation report for Eagle Harbor site, Kitsap Co., WA, by CH_2M Hill for U.S. EPA, Region 10, Seattle, WA, 1991.
33 U.S. EPA, Guidance for data useability in risk assessment (Part A), 9285. 7-09A. Washington, DC, 1992.
34 L.J. Field, A.T. Fritz and C.S. Cairncross, Use of sediment bioassays at Superfund sites in coastal regions, In: Society of Environ. Toxicol. and Chem., 12th Annu. Meeting, Abstract 544, Washington, DC, 1991, p. 136.