United States Environmental Protection Agency
Office of Solid Waste and Emergency Response
Publication 9345.0-051
EPA 540-F-94-012
PB94-963303
September 1994

EPA ECO Update

Office of Emergency and Remedial Response
Hazardous Site Evaluation Division (5204G)

Intermittent Bulletin
Volume 2, Number 1

Using Toxicity Tests in Ecological Risk Assessment

Toxicity tests are used to expose test organisms to a medium—water, sediment, or soil—and evaluate the effects of contamination on the survival, growth, reproduction, behavior and/or other attributes of these organisms. These tests may help to determine whether the contaminant concentrations in a site's media are high enough to cause adverse effects in organisms. Generally, toxicity tests involve collecting samples of media from a site and sending them to a toxicity laboratory, where the tests are performed. On occasion, investigators[1] measure toxicity by exposing test organisms to soil or water on site—these are known as *in situ* tests.

As the general guidelines at the end of this Bulletin indicate, not all sites require toxicity tests. But where they are used, toxicity tests can contribute to ecological risk assessments in specific ways and at different stages in the assessment.

1. *Toxicity tests can demonstrate whether contaminants are* ***bioavailable.***[2] The presence of a contaminant does not of itself indicate a potential for adverse effects. A contaminant can have toxic effects only if it occurs in a bioavailable form. Sometimes the presence of abrasives, such as the talc in pesticides, can damage an organism's body covering, thereby increasing the bioavailability of certain contaminants for that organism.

2. *Toxicity tests can evaluate the aggregate toxic effects of all contaminants in a medium.* Many Superfund sites present a complex array of contaminants, with a mixture of potentially harmful substances present in the media. At such sites, chemical data alone cannot accurately predict the toxicity of the contaminants. Rather, toxicity tests measure the aggregate effects of contaminated media on organisms. These effects result from characteristics of the medium itself (such as hardness and pH, in the case of water), interactions among contaminants, and interactions between contaminants and media. Consequently, observed toxicity test results may often vary from those predicted by chemical data alone.

3. *Toxicity tests can evaluate the toxicity of substances whose biological effects may not have been well characterized.* The contaminants at a Superfund site might include substances that have not been previously investigated regarding their toxicity to wildlife or other organisms. Consequently, the scientific literature contains no relevant data concerning these substances. At such sites, toxicity tests of media samples

[1] The term "investigator" refers to the individual charged with responsibility for designing and/or carrying out any part of an ecological risk assessment. Investigators can include government scientists, contractors, or university scientists. However the site manager (remedial project manager or on-scene coordinator) retains ultimate responsibility for the quality of the ecological risk assessment.

[2] Bioavailability is the presence of a substance in a form that organisms can take up. (Note that specialized terms appear in boldface and are defined either in the text or in accompanying footnotes.)

IN THIS BULLETIN

indicate the combined toxicity of *all* contaminants, including those that have not been previously tested.

4. *Toxicity tests can characterize the nature of a toxic effect.* Investigators can use toxicity tests to learn whether contaminant concentrations have lethal or sublethal effects. Some examples of sublethal effects include reduced growth, impaired reproduction, and behavioral changes.

5. *Toxicity tests can characterize the distribution of toxicity at a site.* An investigator can have toxicity tests performed on samples from a variety of locations at the site. In some instances toxicity tests may be a cost-effective way to determine the spatial extent of toxicity and identify areas with high levels of toxicity.

6. *Toxicity tests can be used to develop remedial goals.* Acceptable levels of toxicity, as measured by toxicity tests, can form a criterion for remedial goals. For example, a goal might be to reduce the toxicity of pond water over a stated time period. The remedial goal would specify the level to which toxicity should be reduced and the species in which toxicity should be measured. The species should be representative of the site and sensitive to its contaminants. The species also should relate to the overall **assessment endpoints**.[3]

7. *Toxicity tests have a role in monitoring.* Toxicity tests can be used to monitor the remediation of a Superfund site. Specifically, toxicity testing can indicate whether sources of contamination have been contained and whether remedial measures are reducing toxicity.

8. *Toxicity tests have a role in determining a site's post-remediation potential to support a viable ecological community.* For example, if a stream or waterbody receives contaminants from numerous sources, including a Superfund site, upstream toxicity testing may help to determine what the water's potential for supporting a viable ecological community might be if the Superfund loadings are removed and the other sources remain unchanged.

Toxicity tests include a broad spectrum of tests, differing in the species and exposure media they use and the effects they measure. In making decisions about whether to conduct toxicity tests, which tests to choose, and how many to perform, investigators are well advised to seek advice from qualified experts, such as those serving on a Regional Biological Technical Assistance Group (BTAG).[4]

This Bulletin first describes two major classes of toxicity tests—acute and chronic—and then explores the elements that an investigator needs to consider in planning toxicity tests. Finally, the Bulletin offers general guidance on when to use toxicity tests and how to select those appropriate to different sites. The companion document, "Catalogue of Standard Toxicity Tests for Ecological Risk Assessment" (*ECO Update* Vol. 2, No. 2), provides an annotated list of standardized tests appropriate for use with different media.

[3] An assessment endpoint is an ecological characteristic that may be adversely affected by site contamination and that, at a Superfund site, can help to drive remedial decision making (U.S. EPA, 1992).

Measurement Endpoints In Toxicity Testing: Acute Vs. Chronic Tests

Toxicity tests can measure lethal and/or sublethal effects. These effects are known as **measurement endpoints**: that is, they are ecological attributes that may be adversely affected by exposure to site contaminants and that are readily measurable. In addition, each measurement endpoint is closely related to an assessment endpoint. Because of this close relationship, a measurement endpoint can approximate or represent the assessment endpoint if the assessment endpoint is not amenable to direct measurement (U.S. EPA, 1992).

Acute toxicity tests are short-term tests that measure the effects of exposure to relatively high concentrations of chemicals. The measurement endpoint generally reflects the extent of lethality.

Chronic toxicity tests, on the other hand, generally are longer-term tests that measure the effects of exposure to relatively lower, less toxic concentrations. For a chronic toxicity test, the measurement endpoint concerns a sublethal effect (e.g., reproduction, growth) or both lethality and a sub-lethal effect.

Acute Toxicity Tests

A typical acute toxicity test exposes test organisms to a series of dilutions of a site's medium and records deaths occurring over a specified period of time, usually 24 to 96 hours. Results can be analyzed by comparing percent mortality of organisms exposed to site media to percent mortality for organisms exposed to uncontaminated media. (See section below entitled "The Reference Site.") Alternatively, results of an acute toxicity test can be analyzed to estimate the dilution of the medium at which 50 percent of the organisms died. This dilution (also referred to as a concentration), called the $\mathbf{LC_{50}}$, is the median lethal concentration. When an acute toxicity test reports an LC_{50}, the test results usually will specify the test duration, the test species, and the life cycle stage of the test species (e.g., the fathead minnow 96 hour LC_{50}). Since LC_{50}s are point estimates, which are estimates of the effects from specific concentrations of contaminants, coefficients of variation can be calculated for them. (See section below entitled "Statistical Analysis.")

[4] These groups are sometimes known by different names, depending on the Region. Readers should check with the appropriate Superfund manager for the name of the BTAG coordinator or other sources of technical assistance in their Region. A more complete description of BTAG structure and function is available in "The Role of BTAGs in Ecological Assessment" (*ECO Update* Vol. 1, No. 1).

With some test organisms, toxicologists find death difficult to determine unequivocally. In tests using such organisms, toxicologists evaluate another effect, such as immobility, that correlates closely with death. As with death for a measurement endpoint, results can be analyzed by comparing percent effect for organisms exposed to site media and those exposed to uncontaminated media. Alternatively, data can be analyzed to estimate the dilution at which 50 percent of the organisms displayed the effect. This dilution (also referred to as a concentration), called the EC_{50}, is the median effective concentration. When an acute toxicity test reports an EC_{50}, the test results will specify the effect, the test duration, the test species, and the life cycle stage of the test species. Like the LC_{50}, the EC_{50} is a point estimate and a coefficient of variation can be calculated for it.

Acute toxicity tests are short-term tests that measure the effects of exposure to relatively high concentrations of chemicals.

Chronic toxicity tests generally are longer-term tests that measure the effects of exposure to relatively lower, less toxic concentrations.

In still other approaches to evaluating results, the laboratory analyzes the data for the **Lowest Observed Effect Concentration (LOEC)**, which is the highest dilution causing statistically significant toxic effects, or the **No Observed Effect Concentration (NOEC)**, which is the lowest dilution at which no statistically significant toxic effects occurred.[5] Statistically determined using hypothesis testing, LOECs and NOECs are not point estimates and consequently coefficients of variation cannot be calculated for them.

Chronic Toxicity Tests

A chronic toxicity test exposes test organisms to a series of dilutions of a site's medium and measures sub-lethal effects, and in some cases lethal effects as well. Sublethal effects may include growth reduction, reproductive impairment, nerve function impairment, lack of motility, behavioral changes, and the development of **terata**, which are structural abnormalities. Results can be analyzed in several ways. One is simply by a direct comparison between percent effect occurring in organisms exposed to site media and those exposed to uncontaminated media. Other approaches to analysis determine the EC_{50}, the LOEC, or the NOEC.

Ecological Significance of Sublethal Effects

Although it would be an oversimplification to extrapolate from the outcome of chronic toxicity tests to ecological conditions at a Superfund site, site managers need to be aware that the sublethal effects that chronic toxicity tests measure in laboratories are ecologically significant effects when they occur in the environment. For example, reduced growth can lead to decreased production, smaller size, lower fecundity (eggs or young per female), increased susceptibility to predation, and other effects. Reproductive impairment can reduce the population size and also bring about changes in a population's age structure. Production of individuals with terata can adversely affect a population because these individuals have a lower growth rate, are generally unable to reproduce, and have an increased susceptibility to predation.

A Comparison of Acute and Chronic Toxicity Tests with Respect to Time, Cost, and Resolution

In general, acute and chronic toxicity tests differ in the amount of time required to perform them, their cost, and their resolution.

- Because chronic tests extend through either a life cycle or a critical developmental phase, they generally require more time to perform than acute tests with the same type of test organisms.

- Requiring more time to complete than acute tests, chronic tests also can require more funds. A chronic test also may require more resources and increased numbers of laboratory analyses, further increasing the cost of the test.

- Chronic tests have greater resolution than acute tests. For example, consider a chronic test that exposes invertebrates to site surface water and records the number of young they produce. In a highly toxic medium, the organisms will die. In a less toxic medium, they may survive, but their reproductive capacity may be impaired when compared with controls maintained in an uncontaminated medium.

Elements in a Toxicity Assessment

The investigator needs to consider many elements when planning a toxicity assessment: the objective, the reference site, the medium analyzed, the test organisms, the test methodology, the level of effort, the test site, and quality assurance/quality control (QA/QC) standards. By the choices that he or she makes, the investigator can tailor the toxicity assessment to meet the needs of the site and its stage in the Superfund process.

[5] As used in this Bulletin, LOEC is synonymous with Lowest Observed Adverse Effect Concentration (LOAEC) and Lowest Observed Adverse Effect Level (LOAEL), and NOEC with No Observed Adverse Effect Concentration (NOAEC) and No Observed Adverse Effect Level (NOAEL).

The Objective

As with any study, before planning a toxicity assessment the investigator needs to set clear objectives. In particular, the assessment's objectives need to include some that address the medium of concern, the characteristics of the contaminants of concern, and the potential **ecological components**.[6] For example, if the study asks whether soil on the site is toxic to macroinvertebrates, then the study will need to analyze bulk soil rather than an **elutriate**[7] and will need to use an appropriate test organism.

The objectives of a toxicity assessment should indicate the level of effort appropriate to the assessment. For example, determining whether a particular medium is toxic would generally require a low level of effort. Such a study might specify only two species of test organisms and undiluted medium collected from a limited number of sampling locations. If the objective of a toxicity assessment is to determine the appropriate range of dilutions for conducting further tests (if these prove necessary) at a highly contaminated site, a higher level of effort would be necessary. Such a study might specify using a series of tenfold dilutions of media collected from locations known to have high contaminant concentrations. A reasonably detailed characterization of a site's toxicity would imply a high level of effort. This type of study might include test organisms at different **trophic levels**[8] (such as an alga, a macroinvertebrate, and a fish), several sampling locations (possibly based on a grid or selected from upstream and downstream areas), and several dilutions of medium.

The Reference Site

When planning a toxicity assessment, an investigator selects a reference site that as closely as possible mirrors the characteristics of the site medium being analyzed but is unaffected by site contamination. Analyzing a sample from the reference site allows the investigator to measure background conditions. The investigator should try to locate the reference site as close as possible to the Superfund site so that the reference site will accurately reflect the site's conditions. Yet the reference site should lie at a great enough distance from the Superfund site to be unaffected by site contamination. Provided that pollutant loading from other sources does not occur upstream, an upstream location may provide an appropriate reference site for a Superfund site with contaminated surface water. Soil type and texture, vegetation, and slope are important considerations in selecting a reference site with the appropriate terrestrial characteristics.

The Medium

Toxicity tests vary as to the media they analyze. Aquatic tests evaluate freshwater, marine, or estuarine samples. A few tests are designed specifically to analyze bulk sediment samples, and a few are specific for bulk soil samples. Bulk sediment or soil tests specifically address toxicity in the test medium. Alternatively, laboratory technicians can prepare elutriates of sediment or soil samples and analyze the elutriates by means of aquatic tests. Toxicity tests using elutriates give information about the transfer of contaminants from sediment or soil to water. Such information is most valuable when predicting effects of runoff or leaching from soil or determining the advisability of remediating a site by dredging contaminated sediments.

A toxicity test also should include measurements of the appropriate physical and chemical parameters of the sample medium. For water, these parameters might include alkalinity, hardness, pH, temperature, dissolved oxygen, total dissolved solids, and total organic carbon. For a sediment sample, grain size, percent water, pH, total organic carbon, and/or other parameters may prove important to know.

A toxicity test should include measurements of the appropriate physical and chemical parameters of the sample medium.

In some cases the physical or chemical parameters of the test medium require adjustment in order to meet the conditions of a test protocol. Sediment or soil may require dewatering. Water samples may need to have their pH, hardness, or dissolved oxygen content adjusted. Such adjustments can change the solubility, bioavailability, or toxic properties of sample constituents and therefore should be avoided or minimized wherever possible. If the test medium requires adjustment, the investigator should allow a portion of it to remain unadjusted. This unadjusted portion is used in a parallel control that will indicate whether the adjustment contributes to, masks, or has no effect on toxicity. In cases where the test medium requires adjustment, the investigator should evaluate the **data quality objectives (DQOs)**[9] to determine whether the adjustments would interfere with the study's objectives.

For many toxicity tests investigators must dilute sample media to determine LC_{50}s, EC_{50}s, LOECs, or NOECs. Protocols for aquatic tests generally specify using specially treated laboratory water as a diluent, but natural water can be used as well. Diluting material for soils or sediments can consist of artificial soil prepared in the laboratory.

[6] An ecological component is an individual organism, a population, a community, a habitat, or an ecosystem that may suffer adverse effects as a result of site contamination.

[7] An elutriate (or eluate) is the solution obtained when water removes substances adsorbed to sediment particles.

[8] A trophic level is a stage in the flow of food from one population to another. For example, as primary producers (organisms that convert the energy from sunlight to chemical energy) plants occupy the first trophic level, and grazing organisms occupy the second trophic level.

[9] Data quality objectives (DQOs) are statements that define the level of uncertainty that investigator is willing to accept in environmental data used to support a remedial decision. DQOs address the purpose and use of data, the resource constraints on data collection, and any calculations based on the data.

Test Organisms

Toxicologists have based their selection of test organisms on several factors: sensitivity to a variety of substances, availability, representativeness of a variety of ecosystems, and ease of maintenance and culture under laboratory conditions. For aquatic tests, the most frequently used test organisms are those employed for toxicity testing for National Pollutant Discharge Elimination System (NPDES) permits. Table 1 summarizes information about the organisms used in the standardized tests, while Figure 1 illustrates a few of these organisms.

When choosing from among the available standard test organisms, the investigator should select a species that is representative of resident organisms, sensitive to site contaminants, relevant to the overall assessment endpoints, and consistent with DQOs. In a toxicity test, the test organisms serve as surrogates for organisms present on the site. For instance, although fathead minnows (*Pimephales promelas*), a common test organism, may not occur on the site, they can serve as surrogates for other fish. Consequently an LC_{50} for fathead minnows can serve as a measurement endpoint for the assessment endpoint "survival of the minnow populations in a specific stream that flows through the site." In a broader context, fathead minnows might represent all warmwater fish on a site, since research has shown that organisms at the same taxonomic level (level of classification, such as genus or family) often respond similarly to a contaminant (Baker, 1989). When selecting test organisms, the investigator should keep the study's DQOs in mind. If the investigator's selection is not consistent with the DQOs, the applicability of the test data to the site is questionable.

Although the existence of well-established protocols and considerable historical data makes the standard test organisms useful, in some cases investigators find that none of the standard organisms is representative of a site's ecosystem. If this situation occurs, the investigator must account for this lack of representativeness when interpreting test results. Alternatively, the investigator may decide to use a "non-standard" or alternative species instead of the one specified in the test protocol. The alternative species might better represent resident organisms, show greater sensitivity to the site's contaminants, or be more consistent with the study's DQOs. State resource agencies can readily provide information on resident species. Using resident species as an alternative species has the potential of providing direct information about the toxic effects to site species. Several criteria must be specified for the use of alternative test species, including the source for the test organism, the age range suitable for the test, a means for eliminating variability in the organism's condition, and conditions suitable for the test. In addition, if the organism must be collected rather than purchased, the investigator will have to establish standards for ensuring accurate identification and also should meet all local, state, and federal requirements concerning the collection of organisms.

When choosing from among the available standard test organisms, the investigator should select a species that is representative of resident organisms, sensitive to site contaminants, relevant to the overall assessment endpoints, and consistent with DQOs.

Generally, using alternative species increases the cost of conducting toxicity tests, especially when the investigator needs to determine optimal conditions for conducting the test. However, the investigator, in consultation with the BTAG, may decide that the added usefulness of the results justifies the extra expense. In such a case, the investigator may be able to reduce the added expense by employing a laboratory experienced in the use of the species selected for the study.

Test Method

Standard toxicity tests can employ a variety of methods for collecting samples and for exposing test organisms to media. Designing a toxicity assessment for a site requires the investigator to select the most appropriate methods for studying the issues for that site.

Field biologists can collect media samples for testing either by the grab or the composite method. As the name implies, a grab sample is a single sample, usually entailing little time and minimal equipment to collect. When the investigator expects the site's contaminant picture to change little over time, a single grab sample per location may adequately represent contamination. A composite sample, on the other hand, is a mixed sample, which may be collected at a single location over a specified period of time or at multiple locations at one time. When sampling a stream with a highly variable flow rate, the investigator can specify the collection of a flow weighted composite sample. The BTAG can advise the investigator as to the preferred collection method for a particular site.

Toxicity tests analyzing water or elutriates of soil or sediment can expose test organisms using the same sample medium throughout the test or arranging for limited replacement of medium. Those using the same sample medium throughout are called **static** tests, while **static-renewal** tests are those that replace all or part of the sample medium at specified times during the test. Since this approach requires little space, manpower, and equipment, static tests are comparatively simple and inexpensive to perform. In addition, static tests require only small sample volumes of one to 20 liters.

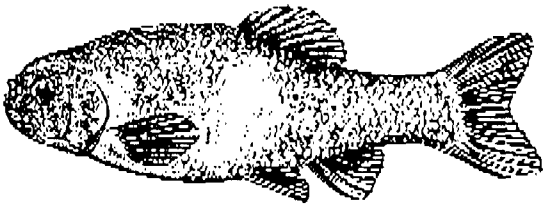
On the other hand, static tests, particularly those without renewal of media, do have certain limitations. Over the course of a non-renewal test, test organisms can deplete the dissolved oxygen in the sample and suffer adverse effects unrelated to toxicity. Alternatively, contaminants can break down, volatilize, or adhere to the walls of the container. As a result, the test might not accurately reflect the medium's toxicity. Finally, as organisms metabolize they release substances, such as carbon dioxide and

Table 1. Plant and Animal Species Used in Standard Toxicity Tests*

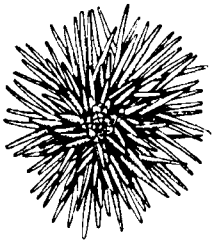
Medium	Test Organism	Test Temp (°C)	Life Stage
FRESHWATER	VERTEBRATES		
	Brook trout *(Salvelinus fontinalis)*	12	30–60 days
	Rainbow trout *(Oncorhynchus mykiss)*	12	15–30 days
	Fathead minnow *(Pimephales promelas)*	20–25	1–14 days
	INVERTEBRATES		
	Amphipod *(Hyalella)*	20 or 25	7–14 days
	Waterflea *(Daphnia magna, Daphnia pulex, Ceriodaphnia)*	20 or 25	1–24 hours
	Mayfly *(Hexagenia limbata, Hexagenia bilineata)*	17, 20–22	Young nymph
	Midge *(Chironomus)*	20 or 25	First to second instar
	ALGA		
	Selenastrum capricornutum	25	4–7 day stock culture
MARINE	VERTEBRATES		
	Sheepshead minnow *(Cyprinodon variegatus)*	20 or 25	1–14 days
and	Silverside *(Menidia* species)	20 or 25	9–14 days
	INVERTEBRATES		
ESTUARINE	Sea urchin *(Arbacia punctulata)*	20	< 1 hour old
	Mysid shrimp *(Mysidopsis)*	20	1-5 days
	ALGA		
WATERS	*Champia parvula*	23	Sexually mature
FRESHWATER SEDIMENT	Amphipod *(Hyalella azteca)*	20–25	7–14 days
	Midge *(Chironomus tentans and Chironomus riparius)*	20 or 25	First to second instar
MARINE SEDIMENT	Amphipod *(Rhepoxynius abronius)*	15	Mature 3–5 mm, mixed sex
	Amphipod *(Eohaustorius estuarius)*	15	Mature 3–5 mm, mixed sex
	Amphipod *(Ampelisca abdita)*	20	Immature, or mature females only
	Amphipod *(Grandidierella japonica)*	15–19	Immature 3–6 mm, no females with embryos
SOIL	Earthworm *(Eisenia foetida)*	22	300-600 mg adult
	Lettuce *(Latuca sativa)*	24	Seed

*Compiled from ASTM, 1992c; Greene et al, 1989; US Army Corps of Engineers, 1993; Weber et al, 1988; Weber et al, 1989; Weber et al, 1991.

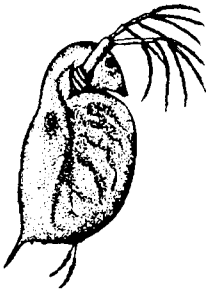
Figure 1. Test Organisms Commonly Used in Toxicity Tests
The organisms shown in these figures are the adult life stages of the test organisms, not necessarily those life stages used in toxicity tests.



Fathead minnow *(Pimephales promelas)*
50 mm.

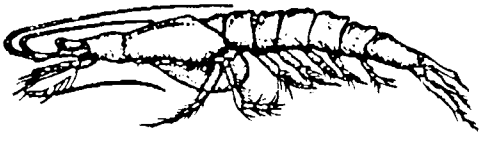


Purple sea urchin
(Stronglyocentrotus purpuratus)
6 to 12 cm

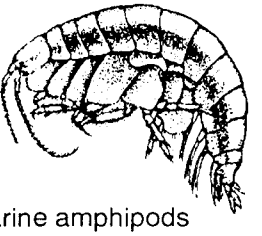


Waterflea *(Daphnia)*
up to 3.5 mm for *Daphnia pulex*

Macroalga *(Champia)*
Actual size of branch tip 25 mm.



Mysid shrimp *(Mysidopsis)*
4.4 to 9.4 mm for *Mysidopsis bahia*



Marine amphipods
Species used in toxicity tests usually
3-6 mm.

Reproduced by courtesy of Aquatic Research Organisms.
P.O.Box 1271, Hampton, NH 03842.

wastes, called metabolites. As these metabolites accumulate, they can prove toxic to the test organisms. Alternatively, they may interact with contaminants and alter the medium's apparent toxicity. The static-renewal design overcomes to a certain extent the disadvantages of the nonrenewal design.

In place of the static design, aquatic toxicity tests may use a **flow-through** method, continuously pumping fresh sample medium through test chambers. With this method, dissolved oxygen remains relatively high and metabolites are flushed away. The flow-through approach also has the advantage of minimizing loss of toxics to degradation, adsorption, and volatilization. When conducted on-site, flow-through tests can provide information about fluctuations in toxicity.

Disadvantages of the flow-through method include its expense and inconvenience. This approach uses complex equipment that requires more maintenance than the equipment used for static tests. Large volumes of sample and diluent are needed for flow-through tests. As a result, these tests are more expensive than equivalent static tests.

Level of Effort

The quantity and nature of information provided by toxicity tests varies considerably with the level of effort mandated by the study's objective. One factor determining the level of effort is the number of species used as test organisms. When the investigator selects a test species that is quite sensitive to the site's contaminants relative to other types of organisms in the community, then a single-species test can suggest the community's maximum susceptibility to the site's contaminants. At a higher level of effort, employing more than one test species may indicate whether single-species tests have underestimated or overestimated the site's toxicity. However, when evaluating water, toxicity tests should always include at least two species, a fish and an invertebrate, unless the site has only one contaminant and either fish or invertebrates are known to be insensitive to that contaminant.

The level of effort also varies with the range of media concentrations analyzed and the duration of the tests.

- **Screening tests**, which may be either acute or chronic tests, evaluate only the undiluted sample. Positive values on screening tests may indicate the need to proceed to definitive tests.

- **Range finding tests** are abbreviated static acute tests that expose test organisms to a broad range of media dilutions for 8 to 20 hours. Such tests identify the dilutions to use for definitive tests.

- **Definitive tests** provide a dose-response curve and reduced variability. Both acute and chronic tests can be conducted as definitive tests. Investigators can use these three types of tests to translate "level of effort" objectives into appropriate toxicity assessments.

Test Site: Laboratory or *in situ*

In general, toxicity tests are conducted by collecting samples from a site and sending them to a laboratory for testing. Laboratory tests offer the advantage of standardized protocols. In addition, laboratories experienced in performing these tests are generally available.

Alternatively, toxicity tests can occur *in situ*, which means "in place." That is, the investigator exposes test organisms to soil or water on the site. *In situ* tests give organisms continuous exposure to the site media under actual environmental conditions such as temperature, stream flow, and light.[10] Consequently, data from *in situ* tests provide a more realistic assessment of toxicity than do data from laboratory tests. *In situ* tests also provide a more direct means of comparing toxicity data with estimates of exposure derived from field data. Ultimately, such comparisons help to characterize the site's ecological risk.

On the other hand, *in situ* tests have certain disadvantages. In particular, the investigator lacks control over the conditions under which an *in situ* test occurs. For example, temperature may vary considerably over the course of an *in situ* test, whereas in the laboratory temperature would be carefully regulated. When interpreting data from *in situ* tests, the investigator should consider how the varying conditions of the test could affect results. Another issue associated with *in situ* tests concerns logistics, which can prove difficult in adverse weather conditions.

At the present time, *in situ* tests generally consist of standard laboratory tests adapted for on-site use. Some commonly used tests employ earthworms or lettuce seeds as standard test organisms for soil toxicity tests. For *in situ* aquatic toxicity tests, fish, clams, and oysters are often used. Tests may instead use alternative or resident species. In practical terms, to conduct *in situ* tests, test vegetation is maintained in plots and test animals are held in containers on site. The earthworm test, in particular, has been successfully adapted for *in situ* use. In fact, some field biologists routinely perform this test *in situ* rather than having personnel conduct it as a laboratory test.

Statistical Analysis

The statistical analysis of toxicity test results depends upon the measurement endpoint. As mentioned earlier, LC_{50}s and EC_{50}s are point estimates; that is, they are estimates of effects from specific dilutions of contaminants. To calculate point estimates, test data are analyzed using regression models that assume the less dilute the sample, the greater will be the effects. Coefficients of variation can be calculated for point estimates.

LOECs and NOECs compare results at test dilutions with controls to determine whether the results are significantly different. LOECs and NOECs are calculated by means of a statistical method called hypothesis testing. Coefficients of variation cannot be calculated for values determined using hypothesis testing. Note too that the values obtained for LOECs and NOECs can vary considerably depending on the specific series of dilutions used in the test.

For detailed information concerning statistical analysis of toxicity test data, investigators can consult Weber et al. (1988, 1989, 1991).

[10] *In situ* tests should not be confused with tests performed in a mobile laboratory brought to the site. A mobile laboratory performs toxicity tests under standard laboratory conditions, not site conditions.

Quality Assurance/Quality Control (QA/QC) Standards

The investigator must specify standards for quality assurance and quality control (QA/QC) to ensure that toxicity tests produce reliable results. QA/QC standards describe appropriate sample handling and collection. Such matters as the container type, storage temperature, and maximum permissible storage time affect the reliability of data. For example, elapsed time and exposure to air can alter sample properties or result in the loss of volatile chemicals. Generally, aquatic samples should be stored at 4°C until used. Although holding time limitations for sediments vary with individual sediment properties and contaminant characteristics, the American Society for Testing and Materials (ASTM) recommends storing sediments at 4°C and using them within two weeks of collection.

Test parameters should specify which controls to perform and what values to accept for controls. For example, in an acute toxicity test with *Daphnia pulex*, a control consists of test organisms exposed to dilution water only. For test results to be considered valid, at least 90% of the animals in the control must survive the test.

Finally, QA/QC measures need to ensure that the data collected can support the appropriate statistical analysis.

The BTAG can advise the investigator as to whether the proposed QA/QC standards are adequate.

General Guidelines for Choosing Toxicity Tests

In an ecological risk assessment of a Superfund site, the investigator must decide whether toxicity testing will contribute to the assessment and, if so, which and how many tests to perform. The great differences among Superfund sites—differences in size, terrain, and contaminant profile, to name a few—make a rigidly standardized approach to toxicity studies unworkable . However, a few widely accepted general guidelines do exist:

- *Do not perform toxicity studies at a site where the contaminants of concern do not cause effects measured by toxicity tests.* For example, polychlorinated biphenyls (PCBs) bring about reproductive effects that many toxicity tests do not detect. The investigator should consult the BTAG to find out whether toxicity tests can detect the type of effects caused by a site's contaminants.

- *At a site where several substances have contaminated surface water, aquatic toxicity testing generally should include both a fish and an invertebrate species.* At some sites, it may prove advisable to include additional test organisms, as well.

- *Select test organisms that are sensitive to the site contaminants.* For example, among standard freshwater test organisms, water fleas of the genus *Ceriodaphnia* and the embryos and larvae of fathead minnows are sensitive to a broad spectrum of contaminants. Of the two species of midges used to conduct standard tests of freshwater sediments, *Chironomus riparius* is the choice when metals are the contaminants of concern. When conducting initial tests of saltwater species, it is usually appropriate to use a grass shrimp, penaeid shrimp, or mysid, because these invertebrates often are more sensitive than fish. As a final example, algae should be used for initial tests when herbicides and materials with suspected phytotoxicity are detected in fresh or salt water.

  These are only a few examples of how test organisms differ in their sensitivity to contaminants. Again, investigators will want to consult the BTAG for assistance in selecting appropriate test organisms.

- *When testing water, select a test organism that can tolerate the water's condition.* For example, some organisms, such as the waterflea *Ceriodaphnia,* are extremely sensitive to water hardness. Also, certain inland waters have high enough salinity to make the use of freshwater test organisms inadvisable.

To extend the limited guidelines offered above, several EPA scientists were interviewed to learn how they design toxicity assessments. Each scientist offered a somewhat different outline and set of priorities. These differences reflect differences in site characteristics and geography, such as degree of urbanization and amount of rainfall. In spite of the great variety in hazardous waste sites, three general designs for toxicity assessments emerged from these interviews. The following section presents the designs. Where a design refers to specific toxicity tests, these are described in the companion Bulletin, "Catalogue of Standard Toxicity Tests for Ecological Risk Assessment" (*ECO Update* Vol. 2, No. 2).

Design 1

Design 1 is based on the premise that designing a meaningful toxicity assessment requires considering a contaminant's mode of action, the level of site contamination, the sensitivity of the test type (acute or chronic), and the sensitivity of the test organism. In particular, sites are first reviewed to determine whether contaminants are expected to cause effects detectable by toxicity tests that meet the stated DQOs for that site. The investigator then matches contaminant levels with test type. For example, in a heavily industrialized site with high levels of contaminants, many of the samples in a chronic test may give highly positive results. Consequently, the study would provide little information about differences in toxicity at various locations on the site. At such a site, acute tests might distinguish the varying toxicity of different locations, thereby identifying the more highly impacted areas and helping to establish priorities for remedial decisions.

Similar concerns are addressed when choosing test organisms: too sensitive a test organism results in a lack of discrimination and too insensitive a test organism can give negative results that are misleading. Also, Design 1 uses screening tests sparingly, because these tests may suffer from over-interpretation of results.

Design 2

Design 2 focuses on planning toxicity assessments for sites where chemical data do not clearly indicate whether the site contaminants represent an ecological problem requiring further action. The premise is that sites with intermediate levels of contamination will most likely require toxicity testing. Where little

contamination has occurred, the chemical data will generally indicate that the site is not an ecological problem and requires no further ecological risk assessment. A heavily contaminated site will almost certainly require further action.

With this design, the investigator decides on ecological components and endpoints as early as possible, making use of all the information available about the site and its contaminant history. The ecological components and endpoints then serve to guide choices of both toxicity tests and field studies.

If the chemical data indicate the need for toxicity tests, as reflected by the literature, the investigator next decides which media to test. Contamination may have migrated from the initial contaminated medium into other site-associated media. Toxicity test selection occurs next, with the choice between acute and chronic tests depending on the objectives. In the long run, gauging toxicity based on the sub-lethal effects measured by chronic tests proves more protective. In addition, chronic tests are considered more appropriate if an organism spends most of its time on-site. Conversely, acute tests better mimic conditions for organisms that spend a limited amount of time on-site, such as migratory animals or those with a large home range. Acute tests also prove useful when the investigator has concerns about significant differences in toxicity at different locations on the site. If initial tests show no toxic "hot spots," then further testing may not be necessary.

Design 2 puts toxicity testing in perspective by viewing the ecological risk assessment of a Superfund site as a triad consisting of chemical testing, toxicity testing, and field study:

- Chemical testing indicates the presence of contamination.
- Toxicity tests then explore whether biological effects are possible.
- Field studies investigate whether actual harm has occurred at the site.

At a highly contaminated site, each leg of the triad will most likely give evidence of impact. At many sites, however, results are not clear-cut. For example, toxicity results might not correlate well with chemical data. Alternatively, field studies might not demonstrate adverse ecological effects. Such sites require careful professional judgment to make a decision regarding ecological effects and the need for remediation.

Design 3

Design 3 begins with a site reconnaissance visit, followed by a "desktop assessment" to determine whether a site requires an ecological risk assessment. The desktop assessment considers the site's background from scoping and also its contaminants, their environmental concentrations, their physical and chemical properties, and the nature of the surrounding area. A site located in an urban industrial area, for example, may not require an ecological risk assessment: the site itself may have no ecological components of concern, and the contaminants may not have a means of migrating to areas having potential ecological components. On the other hand, the existence of a conduit—a stream, a drainage ditch, ground-water gradient, or land grade—that could carry contaminants from the site to surface water, wetlands, or terrestrial habitats would indicate the need for an ecological risk assessment. In this design, toxicity studies are viewed as useful tools and an effective use of Superfund resources at sites requiring ecological risk assessments.

Selection of sampling locations is one of the first tasks undertaken when designing a toxicity study using this design. Especially at large sites, the investigator avoids random sampling, which would generate an unmanageably large number of samples to analyze. Instead, site terrain and contaminant history are carefully studied in order to place sampling points in areas that will extend the knowledge of the site. As a general rule, at a site with contaminated surface water, the sampling plan should specify sufficient samples to characterize fully the potential ecological effects.

When testing samples of surface water, the investigator selects an initial battery of screening tests from among the standardized acute tests used in the NPDES permitting program. These initial tests usually include organisms at three trophic levels (e.g., an alga, an invertebrate, and a fish). If screening level tests show that surface water is toxic, further tests include surface water dilutions and additional test species to characterize the site's toxicity more fully.

This design favors performing initial soil or sediment evaluations at the screening level using either bulk samples, laboratory-prepared elutriates, or **pore water**, the water located between particles and obtained by either centrifugation or filtration. Like surface water samples, elutriates and pore water samples are analyzed using screening level NPDES aquatic tests. If these tests give positive results, the investigator reviews the chemical data and contaminant history for the site before making the serious commitment of resources that testing of bulk soil or sediment entails.

Conclusion

The differences in these three designs largely reflect the fact that an ecological risk assessment of a Superfund site needs to address the site's characteristics and contaminant history. As scientists gain more experience in conducting toxicity assessments, the several designs existing today may evolve toward a common blueprint that will readily accommodate site differences. In addition to reflecting site differences, the designs also reflect strategic differences in the deployment of resources, time, and

As scientists gain more experience in conducting toxicity assessments, the several designs existing today may evolve toward a common blueprint that will readily accommodate site differences.

personnel. Despite these differences, however, the designs have the same end: to assess the toxicity of contaminated media from Superfund sites.

To learn about other designs for toxicity assessments, investigators may wish to consult the scientists in the EPA Environmental Services Division (ESD) in their Region. In making decisions about toxicity testing at a specific site, an investigator should consult with the Regional BTAG. The BTAG will be able to tell the investigator whether the Region has a standard design for toxicity assessments. If the Region does not, the BTAG can advise the investigator whether one of the above designs, one offered by the ESD, or a modification of any of these will further the ecological risk assessment of a particular site. Alternatively, the BTAG may suggest another design that is better suited to a particular site.

References

American Society for Testing and Materials (ASTM). 1992a. *Annual Book of ASTM Standards: Water and Environmental Technology*, Vol. 11.04. American Society for Testing and Materials, Philadelphia, PA.

American Society for Testing and Materials (ASTM). 1992b. *Standard Guide for Conducting Sediment Toxicity Tests with Freshwater Invertebrates*. American Society for Testing and Materials, Philadelphia, PA.

American Society for Testing and Materials (ASTM). 1992c. *Standard Guide for Conducting 10-day Static Sediment Toxicity Tests with Marine and Estuarine Amphipods*. American Society for Testing and Materials, Philadelphia, PA.

Baker, J.P. 1989. "Assessment Strategies and Approaches" in Warren-Hicks, W., B.R. Parkhurst, and S.S. Baker Jr., eds. *Ecological Assessment of Hazardous Waste Sites: A Field and Laboratory Reference*. EPA/600/3-89/013. Environmental Research Laboratory, Office of Research and Development, U.S. Environmental Protection Agency, Corvallis, OR.

Bitton, G., B.J. Dutka, and C.W. Hendricks. 1989. "Microbial Toxicity Tests" in Warren-Hicks, W., B.R. Parkhurst, and S.S. Baker Jr., eds. *Ecological Assessment of Hazardous Waste Sites: A Field and Laboratory Reference*. EPA/600/3-89/013. Environmental Research Laboratory, Office of Research and Development, U.S. Environmental Protection Agency, Corvallis, OR.

Greene, J.C., C.L. Bartels, W.J. Warren-Hicks, B.R. Parkhurst, G.L. Linder, S.A. Peterson, and W.E. Miller. 1989. *Protocols for Short-Term Toxicity Screening of Hazardous Waste Sites*. EPA/600/3-88/029. Environmental Research Laboratory, Office of Research and Development, U.S. Environmental Protection Agency, Corvallis, OR.

Linder, G. et al. 1992. *Evaluation of Terrestrial Indicators for Use in Ecological Assessments at Hazardous Waste Sites*. EPA/600/R-92/183. Environmental Research Laboratory, Office of Research and Development, U.S. Environmental Protection Agency, Corvallis, OR.

U.S. Army Corps of Engineers. 1993. *Evaluation of Dredged Material Proposed for Discharge in Inland and Near Coastal Waters — Testing Manual (Draft)*. Office of Water, Washington, DC.

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

Weber, C.I. 1991. *Methods for Measuring the Acute Toxicity of Effluents and Receiving Waters to Freshwater and Marine Organisms*. 4th edition. EPA/600/4-90/027. Environmental Monitoring Systems Laboratory, Office of Research and Development, U.S. Environmental Protection Agency, Cincinnati, OH.

Weber, C.I., W.I. Horning, D.J. Claim, T.W. Neiheisel, P.A. Lewis, E.L Robinson, J. Menkedick, and F. Kessler. 1988. *Short-term Methods for Estimating the Chronic Toxicity of Effluents and Receiving Waters to Marine and Estuarine Organisms*. EPA/600/4-87/028. Environmental Monitoring and Support Laboratory, Office of Research and Development, U.S. Environmental Protection Agency, Cincinnati, OH.

Weber, C.I., W.H. Peltier, T.J. Norberg-King, W.B. Horning II, F. Kessler, and J. Menkedick. 1989. *Short-term Methods for Estimating the Chronic Toxicity of Effluents and Receiving Waters to Freshwater Organisms*. 2nd edition. EPA/600/4-89/001. Environmental Monitoring Systems Laboratory, Office of Research and Development, U.S. Environmental Protection Agency, Cincinnati, OH.