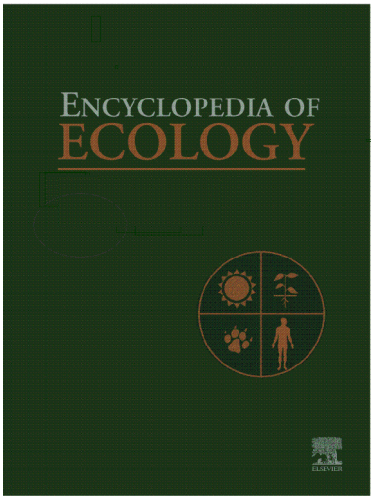




M C Newman and Y Zhao. Ecotoxicology Nomenclature: LC, LD, LOC, LOEC, MAC. In Sven Erik Jørgensen and Brian D. Fath (Editor-in-Chief), Ecotoxicology. Vol. [2] of Encyclopedia of Ecology, 5 vols. pp. [1187-1193] Oxford: Elsevier.

# Ecotoxicology Nomenclature: LC, LD, LOC, LOEC, MAC

**M C Newman and Y Zhao**, Virginia Institute of Marine Science, Gloucester Point, VA, USA





## Introduction

Ecotoxicology is a relatively new science concerned with contaminants in the biosphere and their effects on constituents of the biosphere, including humans. Although relevant effects range from the molecular to the biospheric levels of biological organization, most measures of effect generated by ecotoxicologists are designed to infer adverse effect at the level of the individual organism. This is a consequence of core methods adoption from classic toxicology, a discipline appropriately concerned with effects on individuals. This has created a bias in the ecotoxicology literature that can compromise inferences about effects at higher levels of biological organization such as the population, ecological community, or ecosystem levels. Despite this bias, the methods described herein can and are used pragmatically to infer effects including those occurring at higher levels.

Effect metrics are derived for different kinds of exposures, most notably acute exposures to high concentrations and chronic exposures to low concentrations. Acute exposures are defined in various ways but all definitions reflect a relatively brief and intense exposure scenario. By recent convention, chronic exposures are defined as those exceeding 10% of an individual's life span. However, the definition of chronic exposure varies in literature and depends on the specific methodologies used to produce an effect metric. As an example, some chronic toxicity tests for aquatic species use an exposure duration of 28 days regardless of the longevity of the test species.

## Regression-Derived Effect Metrics

The two general approaches, regression-based and hypothesis testing-based methods, to quantify adverse effects were established early in the history of ecotoxicology. Most regression-based methods are intended to predict the intensity of effect associated with exposure to a toxicant concentration for a specified duration although regression models incorporating exposure concentration and duration simultaneously are becoming increasingly more common. Hypothesis testing-based methods were originally designed to generate some measure of effect in situations in which a regression model cannot be fit with sufficient goodness of fit. However, they have gradually become the effect metrics of choice for chronic exposures that commonly, but not always, produce data less amenable to regression modeling than do acute exposure studies.

Regression-based metrics of effect are generated with well-established test designs. The most widely applied design includes a series of exposure concentrations or doses. There are tests, notably effluent toxicity tests, with slightly different treatments. In the case of an effluent test, the effluent is mixed with different amounts of clean water to generate a series of effluent dilutions. The diluent water is either taken from the receiving water or standard synthetic water is used. The treatment intensity is expressed as (effluent volume)/(effluent volume + dilution water volume) × 100%. Regardless, subsets of test individuals are either exposed to constant levels of toxicant in their surrounding media or food, or given a specific dose of toxicant, perhaps by injection, topical application, or gavage. In the case of a lethal effect, the number of individuals affected at each concentration or dose treatment is tallied after a specified duration (**Figure 1**). The data pairs (concentration, dose, or dilution vs. intensity of effect) generated for the series of concentration, dose, or dilution treatments are then fit to one of several candidate models using conventional regression methods. Most of the commonly applied models are sigmoid functions that accommodate lowest and highest possible limits for effect intensity (**Figure 1**). For lethal effects, the lowest level might be zero or some baseline level of natural ('spontaneous') mortality, and the upper limit of effect is often 100% mortality for the exposed group of individuals.

The most commonly fit sigmoid function is the log normal (probit or normit) model although others such as the log logistic (logit), Weibull (Weibit), or Gompertz (Gompit) often provide excellent fit to these types of data. The conventional probit and probit with spontaneous mortality are the following:

$$p = \Phi(a + b(\ln \text{ dose})) \quad [1]$$

$$p = S + (1 - S)(\Phi(a + b(\ln \text{ dose}))) \quad [2]$$

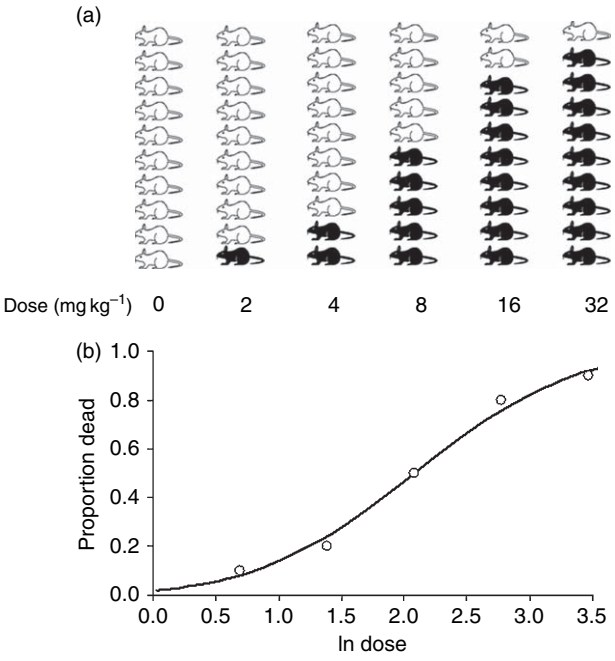


**Figure 1** (a) A typical experimental design used to generate regression-based metrics of effect and (b) the associated sigmoid dose–response curve . The organisms that died in response to the treatment are denoted in black and those still living are denoted in white.

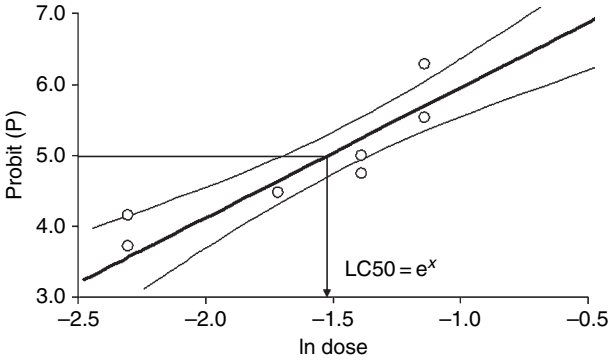


**Figure 2** The probits of the predicted and observed proportions vs. the natural log of the effluent dilution, and the predicted 95% confidence intervals for the data from **Table 1**. The LC50 is estimated by exponentiating the $x$ values corresponding to the probit for 50% mortality.

where $p$ = the probability of or proportion of exposed individuals dying, $S$ = the probability of or proportion of unexposed individuals dying, $\Phi$ = normal cumulative function, $a$ = an estimated regression intercept, and $b$ = estimated regression parameter accounting for the influence of ln dose on $p$. Spontaneous mortality may be included in the model because laboratory culturing conditions are such that some unavoidable mortality occurs or because the longevity of the organism is short relative to the length of the test, natural mortality is to be expected during the test. In either case, the assumption is made that the spontaneous mortality does not influence the relationship between dose/concentration and associated mortality. This may not be an acceptable assumption in some cases.

Such regression models were initially used by laboratory toxicologists to estimate threshold doses below which no effect was expected. However, because the error associated with such estimates was very large, doses predicted to produce certain $p$'s eventually became the norm. Because the prediction error tends to be lowest toward the center of the predicted curve (**Figure 2**), prediction of the concentration producing 50% effect ($p = 0.50$) became the conventional effects metric in mammalian toxicology and was adopted by early ecotoxicologists. Another advantage of prediction for this proportion is that the effects metrics derived by the most common functions (probit and logit) produce very similar results at 0.50. The median lethal dose (LD50) and median lethal concentration (LC50) predicted after a specified duration of exposure are currently the primary metrics of acute lethal effects. For nonlethal effects fit by regression to concentration– or dose–effect models, the median effective dose (ED50) or concentration (EC50) are predicted instead. Despite this convention of predicting median effect levels, a trend has begun that draws the focus of effect assessments more and more often toward lower levels of mortality or effect, that is, LD$x$ or LC$x$ where $x < 50$. We anticipate that this trend will continue, resulting in some changes to the methods described below for predicting LD$x$ and LC$x$ values and their confidence limits. As the emphasis in ecotoxicology shifts downward on the dose/concentration–response curve, more attention will be required in selection of the best among the candidate sigmoid models. Effect metrics are similar for the models at the center of the curve but predictions from the most commonly applied models (e.g., probit, logit, or Weibull) diverge as one makes predictions toward the tails of the distributions.

Predictions of LD50 or LC50 using data from dose/concentration–effect experimental designs can be made using a variety of means. The parametric methods shown in **Figure 3** are most often executed by maximum likelihood estimation (MLE) because it is common to have 0.00 and 1.00 proportions responding (i.e., all individuals in a treatment survived or died) in the data set but the sigmoid functions to which these data are applied never reach 0.00 or 1.00. Which method is applied depends on the data qualities. A model that incorporates spontaneous or natural mortality might be adopted if such baseline mortality is obvious in the data set. In many software packages, the natural mortality can be either specified by the modeler or estimated using various methods by the software. The best sigmoid model fitting the data set can be selected using a goodness-of-fit statistic such as a

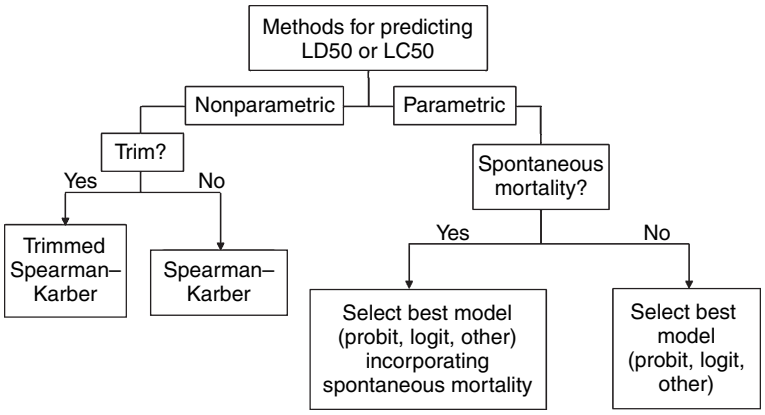


**Figure 3** Methods for predicting LD50 or LC50 values using data from dose or concentration–effect experimental designs.

$\chi^2$ statistic. The advantage of these models is the ease with which estimates and associated confidence intervals can be generated for different $p$ values. A nonparametric approach can be applied instead if the data set does not fit any model acceptably or if the iterative MLE method fails to converge on an acceptable solution. The Spearman–Karber technique with or without trimming of data from the distribution tails is the most commonly applied technique in such a case to generate a LD50 or LC50 estimate and the associated 95% confidence limits.

Data from a mysid shrimp experiment can be used to illustrate the regression method for predicting an LC50 value. Juvenile mysid shrimp were randomly assigned to a series of diluted refinery effluent solutions with ten shrimps per tank. There were seven treatments including a control with duplicate tanks per treatment. Concentrations were expressed as the percentages of total exposure volume made up of the effluent. The mortalities were checked after 48 h of exposure (**Table 1**). These data were fit to candidate models of log normal ('probit') and log logistic ('logit') by an interative maximum likelihood estimation. The associated Pearson $\chi^2$ statistics ($\chi^2 = 4.81$ for log normal, $\chi^2 = 5.01$ for log logistic) indicated that the log normal model provided a better fit than the log logistic. **Figure 2** shows the probits of the predicted and observed proportions dying at 48 h (expressed in probit units) versus the natural log of effluent dilution, and the predicted 95% confidence intervals. (The observed data for the 0% and 100% mortalities were not shown because they do not have corresponding probit values.) Thus the LC50 and the associated lower and upper 95% confidence intervals can be estimated by exponentiating the $x$ values corresponding to the probit of 50% mortality (5), which are 21.9%, 18.4%, and 25.8%, respectively.

The original use of LD50 and LC50 estimates in classic toxicology was as a measure of toxicity. For example, a mammalian toxicologist might use a set of LD50 values to determine the relative toxicities of a series of poisons or to assess how different factors influence the toxicity of a single poison or drug. In such applications, the exposure durations would be set for convenience, for example, acute toxicity after 96 h of exposure because a 96 h test fits conveniently in a workweek, and still generates a meaningful metric of toxicity. So, a $p$ of 0.5 and 96 h test duration might be used for statistical and logistical convenience, not because they reflect pivotal values relative to an acceptable or unacceptable effect to humans.

These regression-derived effect metrics were borrowed by ecotoxicologists who then attempted to apply them to making decisions about the concentration or amount of a chemical that should be a concern if present in an

**Table 1** Dose–response data of an acute toxicity test exposing juvenile mysid shrimp to a simulated refinery effluent

| *Concentration (% effluent)* | *Replicate* | *Number exposed* | *Number dead* |
|---|---|---|---|
| Control | 1 | 10 | 0 |
| | 2 | 10 | 0 |
| 10% | 1 | 10 | 2 |
| | 2 | 10 | 1 |
| 18% | 1 | 10 | 3 |
| | 2 | 10 | 3 |
| 25% | 1 | 10 | 5 |
| | 2 | 10 | 4 |
| 32% | 1 | 10 | 9 |
| | 2 | 10 | 7 |
| 56% | 1 | 10 | 10 |
| | 2 | 10 | 10 |
| 100% | 1 | 10 | 10 |
| | 2 | 10 | 10 |

Modified from table 2 of Buikema AL, Niederlehner BR, and Cairns J (1982) Biological monitoring, Part IV-toxicity testing. *Water Research* 16: 239–262.

environmental media. Given the multiple levels of biological organization that an ecotoxicologist must consider in such a decision, it should be no surprise that these effect metrics do not provide all of the information needed to make an informed decision. Often the duration selected for LD$x$ or LC$x$ estimation is different from that of interest; so extrapolation is required to predict the $p$ associated with an exposure duration other than that used in the test. Such extrapolation can generate unacceptable, or minimally, undefined uncertainties in predictions. The associated uncertainty can be reduced by noting the proportions responding in a series of durations during the test and estimating several LD$x$/LC$x$ for these different durations or by applying survival time regression models instead. Another shortcoming of these effect metrics is that mortality occurring during the period following exposure is rarely measured. Some toxicant exposures produce considerable post-exposure mortality that is important to consider when making predictions of effects to populations exposed in the environment to the chemical of interest. A third shortcoming is not as much one of the regression-related metrics but rather of the decision-making process using these metrics. Often the responsible risk assessors or decision makers lack the expertise or information to determine what level of predicted effect ($p$) should be used as a cutoff for unacceptable adverse effect to exposed individuals, populations, or ecological communities. This can make the regression-related metrics less appealing to assessors and managers than the hypothesis test-based methods described in the following section.

## Effect Metrics Derived with Hypothesis Tests

Developed initially to cope with dose/concentration–response data for which an acceptable model could not be developed, hypothesis test-based methods now are applied heavily in tests of chronic or subtle effects. As will be shown, the intent is to estimate a threshold concentration or dose above which an observable effect might be expected. Most, but not all, relevant statistical methods are conventional hypothesis tests.

The general approach (**Figure 4**) is similar to that shown in **Figure 1** but the variance within and among treatments are assessed instead of developing a dose/concentration–effect model. A series of dose, dilution, or concentration treatments are established with replication within each treatment. After a specific duration, the level of effect manifesting within each treatment is scored and that for each treatment compared statistically to that in the reference treatment. As shown in **Figure 4**, each treatment for which the effect is statistically significantly different from that of the reference treatment is identified (denoted with an asterisk in the figure). The lowest treatment concentration with a response statistically different

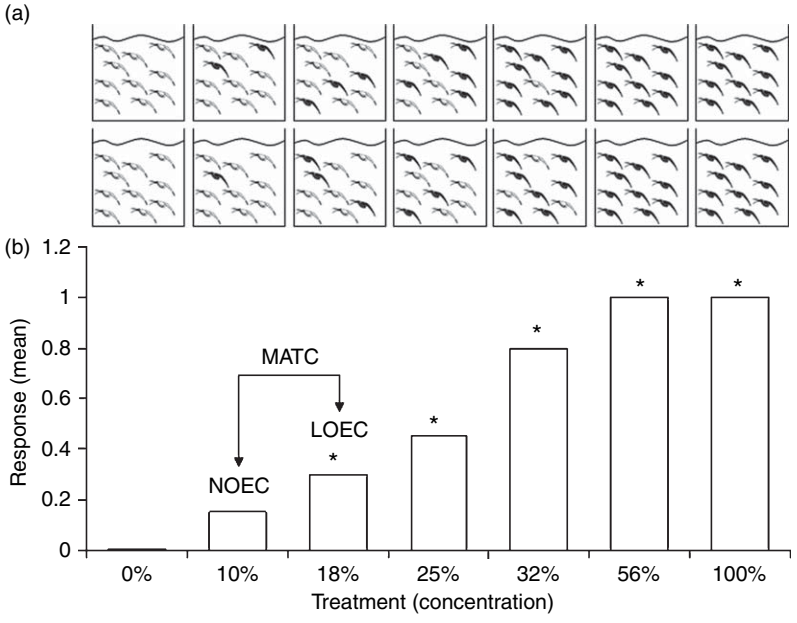


**Figure 4** (a) The experimental design of hypothesis testing-based methods and (b) the determination of NOEC, LOEC, and MATC (data from **Table 1**). The organisms that died in response to the treatment are denoted in black and those still living are denoted in white. The treatments for which the effect is statistically different ($\alpha = 0.05$) from that of the reference treatment are denoted with an asterisk. In this example, the effect was death after the exposure although other responses can, and often are, used in these kinds of experiments.

from that of the reference (e.g., 0%) is called the 'lowest observed effect concentration' (LOEC). The highest treatment concentration with a response that is not significantly different from the reference response is called the 'no observed effect concentration' (NOEC). Although formally a dubious inference from hypothesis testing, the NOEC and LOEC are pragmatically treated in ecotoxicology as the lower and upper bounds for the 'maximum acceptable toxicant concentration' (MATC), that is, a threshold concentration presumed to be 'safe'. Extending this pragmatic approach, the geometric mean of the NOEC and LOEC is sometimes used as the best estimate of the MATC. Considerable debate continues about the acceptability of such interpretations of these hypothesis test-derived metrics.

A range of hypothesis tests are commonly applied to NOEC and LOEC estimation including parametric and nonparametric tests (**Figure 5**). These tests differ in their underlying assumptions and consequent ability to detect a significant difference if there was one, that is, their statistical power. The tests carrying the most assumptions are generally the most powerful. However, the differences in power can be trivial or critical depending on the specific tests being compared and the qualities of the data. As important examples in **Figure 5**, the parametric tests are generally more powerful than the nonparametric tests and tests assuming a monotonic trend with treatment concentration are more powerful than those that do not assume a monotonic trend. With the hypothesis testing approach, the data (concentration, dilution, or dose vs. effect level for each treatment replicate) might be used directly, or as commonly done for proportions responding, transformed in order to meet assumptions of the subsequent hypothesis tests. Formally, the parametric methods can be applied if the data show no evidence of non-normality or heterogeneity of variances among treatments. A powerful parametric trend test (Williams's) can be used if an additional assumption of a monotonic trend (increase or decrease in response) with increasing concentration or dose is justifiable. In some cases such as in the presence of hormesis, a monotonic trend would not be expected. If the assumptions allowing use of the parametric tests are not met, the less powerful nonparametric methods can still be used. If a trend is assumed, then the Jonckheere–Terpstra test can be applied. If not, the less powerful Wilcoxon rank sum test with a Bonferroni adjustment of experiment-wise error rate or the Steel's many-to-one rank test can be used. These last two tests tend to be the least powerful of the hypothesis tests described to this point because they carry the fewest assumptions.

The formal assumptions of and hypotheses tested by these methods differ in important ways. The most important assumption to be met for all is that individuals be randomly assigned to treatments. The results of the hypothesis tests are questionable if this fundamental assumption is not met. The parametric tests further require that the data be normally distributed although most are robust to moderate violations of this assumption. The normality is often tested with a statistic such as the Shapiro–Wilk statistic ($W$). A small value of $W$ (or a $p$ value less than a predetermined $\alpha$ level such as 0.05) leads to the rejection of the null hypothesis of normality. Because the statistical power of the test increases with sample size, when sample sizes are small, a higher $\alpha$ level

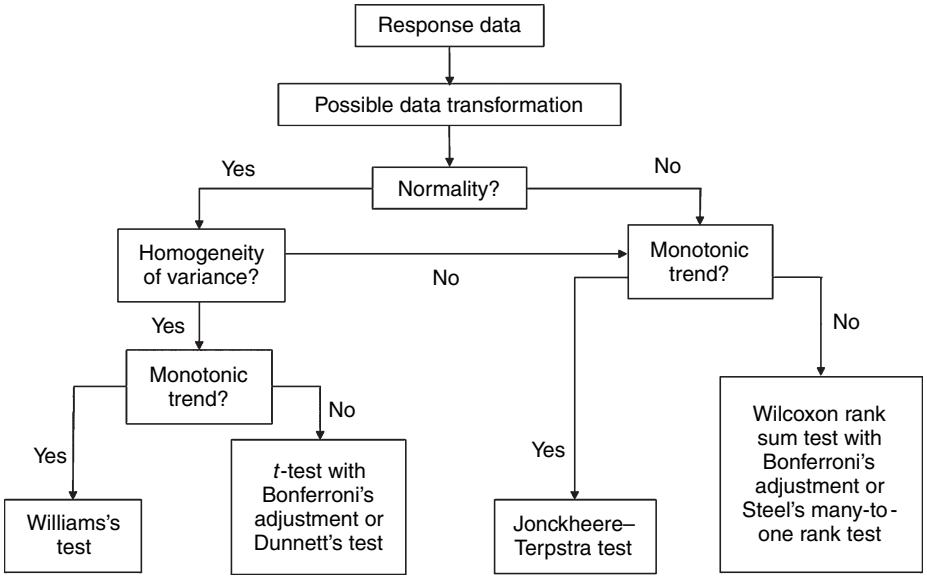


**Figure 5** Parametric and nonparametric hypothesis tests that are commonly applied for NOEC and LOEC estimation.

may be applied in tests of normality. These methods also require that the treatments have the same variances, that is, homogeneity of variances, although again, the methods are robust to moderate deviations from the homogeneity of variances assumption. This assumption can be formally tested with Bartlett's, Levene's, or one of several similar tests. Caution should be taken with the commonly applied Bartlett's test because it can be inaccurate even if the data deviate slightly from being normally distributed.

The common parametric tests differ slightly relative to the exact hypothesis they test. The hypothesis assessed by the *t*-test with Bonferroni adjustment of experiment-wise error rates and Dunnett's test is simply that the mean responses of the treatments are not significantly different from the mean of the reference (control) treatment. Williams's test carries an additional assumption of a monotonic trend (consistently increasing or decreasing effect) with dose/concentration. It tests the null hypothesis that there is no monotonic trend.

The nonparametric methods do not require data normality or homogeneity of variances. With the Wilcoxon rank sum test with Bonferroni adjustment of experiment-wise error rates or Steel's many-to-one rank sum test, the null hypothesis is that observations in the treatments come from the same population. The Jonckheere-Terpstra test is the nonparametric equivalent of the Williams's test in that it has an alternate hypothesis of a monotonic trend. Formally, the null hypothesis for this test is no different in the distribution of responses among the treatments.

The mysid shrimp data can be used again to illustrate the hypothesis testing method (**Figure 4**), although normally more replicates would be recommended. After testing for normality and homogeneity of variance, the data without transformation are tested with Dunnett's one-tailed *t*-test, with the null hypothesis being that the mean response of each treatment is not significantly higher than the control mean (experiment-wise $\alpha = 0.05$). The results show that 10% effluent is the highest concentration whose response is not significantly higher than the control, and 18% effluent is the lowest concentration with the response significantly higher than the control; the NOEC and LOEC were determined to be 10% and 18%, respectively. Accordingly, the MATC could be estimated as the geometric mean of the NOEC and LOEC, or 13.4%. If the log normal (probit) model generated previously had been used to estimate the proportion dead at the NOEC level (10% effluent), the prediction would be 8.0% mortality at the NOEC. The results generally agree between the regression and hypothesis testing approaches although such is not always true.

One shortcoming of the approach of applying these various methods to produce NOEC and LOEC values has already been discussed in the preceding paragraphs – the NOEC and LOEC values can vary for the same data set as a function of the chosen hypothesis test. Also, the NOEC and LOEC values depend heavily on the experimental design and statistical aspects of the calculations that influence statistical power. The power of any test will depend on the number of observations per treatment, number of treatments, and variability in the background response. Literature surveys have demonstrated that the designs normally applied in effects testing have sufficient power to detect an approximately 5–10% effect difference in mammalian toxicology studies and 10–34% effect difference in ecotoxicology studies. But effects less than these levels can have unacceptable consequences. A final shortcoming is that these hypothesis testing methods were not initially designed to infer a biological threshold concentration or dose. A threshold estimated from a test of statistically significant difference is not necessarily a good estimate of a significant biological effect threshold.

## Inferring Consequences of Exposure

Inferring consequences from these effect metrics is challenging but essential. Typical for human effects studies and increasingly common for ecotoxicological studies, the NOEC can be adjusted in a conservative manner to estimate a reference dose (RfD), reference concentration (RfC), or acceptable daily intake (ADI). A common set of adjustments is the following for chronic human exposure:

$$\mathrm{RfD} = \frac{\mathrm{NOEC}}{\mathrm{UF1} * \mathrm{UF2} * \mathrm{UF3} * \mathrm{UF4} * \mathrm{MF}} \quad [3]$$

where UF1 = uncertainty adjustment accounting for variation in natural sensitivity within human populations, UF2 = uncertainty adjustment if extrapolation was performed from animal data to effects to humans, UF3 = uncertainty adjustment if the NOEC comes from a subchronic test data set, UF4 = uncertainty adjustment if the LOEC is used in the calculation instead of an NOEC, and MF = an additional adjustment based on professional judgment. The value for any of these factors can range from 1 to 10 depending on the associated level of uncertainty. This or a similar equation is used to estimate an RfD in the following manner. The literature is searched for all relevant NOEC/LOEC data for the toxicant of interest. The study with the lowest relevant LOEC is identified and the associated NOEC used in the calculation (UF4 = 1). If only the LOEC is available, then the LOEC is applied instead with an UF4 = 10. In the case of chronic human exposure, the RfD is used to estimate the dose level thought to be below that which will cause an adverse effect during chronic exposure. Several types of RfDs are relevant to environmental exposures including short term, subchronic, chronic, or developmental RfDs. The RfD or RfC values may also be developed for different routes of exposure.

With sufficient knowledge, dose- or concentration-effect models can also be applied to estimation of RfD or RfC values. The benchmark dose (BMD) approach uses regression model predictions for a specified effect level (benchmark response) instead of an NOEC or LOEC to estimate the RfD or RfC. Often the lower 95% confidence limit for the estimated BMD (BMDL) is used instead of the NOEC in the above equation to estimate a BMD-based RfD. The UF and MF values can be the same or lower than those used for the NOEC-based approach. Taking the mysid shrimp data as an example, the $BMD_{10}$ could be used to estimate a certain RfC. The $BMD_{10}$ is predicted to be 7.2% effluent. This is the predicted lower 95% confidence interval of the LC10 (10.9% effluent) generated from the log normal (probit) model. As another example, the Environmental Protection Agency (EPA) applied such a BMD approach to determine a human chronic oral methylmercury exposure RfD. Using information from several epidemiological studies, the BMD associated with the lowest 5% of methylmercury-exposed children ($BMD_5$) was chosen as the basis for calculation of the RfD. The primary advantage of this BMD-based approach is that it avoids many of the shortcomings described earlier for the hypothesis test-related effect metrics.

Regardless of how it is calculated, an RfD is used with information about exposure (e.g., inhalation rates, ingestion rates, bioavailability, and exposure duration) to calculate a maximum allowable concentration (MAC, maximum permitted concentration in a particular source such as food, air, drinking water, or soil) or level of concern (LOC, the concentration in the relevant medium above which an adverse effect could manifest).

Protection of human health is facilitated with a set of RfD or RfC values for various exposure scenarios such as acute, prolonged, lifetime, or developmental exposure. Relative to ecotoxicological testing, calculations associated with estimating 'safe' or acceptable exposures are not as straightforward, requiring consideration of consequences at different stages of life cycles of many species and several levels of biological organization. Partial and complete life cycle tests have emerged to address this requirement. A series of tests are conducted at each major life stage of a species, quantifying important effects notionally linked to an individual organism's fitness. The lowest effect metric for the various tests in such a complete life cycle test is used to generate regulatory limits or goals. The cost and difficulty of performing a complete life cycle test has given rise to a less inclusive set of tests (partial life cycle tests) that assess only the life stages thought to be most sensitive. Often these are the early life stages, leading to a battery of tests called early life stage tests. The emphasis during the interpretation of partial or complete life cycle tests is on protection of the individual; however, the EPA stresses the importance of considering population protection for most nonendangered or nonthreatened species existing in ecological communities. That the conventional interpretation of life cycle test-generated effect metrics does not directly address population or community level consequences of exposure is seen as a significant shortcoming in this approach as currently practiced in ecotoxicology. Fortunately, resolving this incongruity between metrics generated with current ecotoxicity tests and prediction of population- and community-level consequences is currently a very active area of research.

*See also:* Acute and Chronic Toxicity.

## Further Reading


Bliss CI (1937) The calculation of the time-mortality curve. *Annals of Applied Biology* 24: 815–852.

Buikema AL, Niederlehner BR, and Cairns J (1982) Biological monitoring, Part IV-toxicity testing. *Water Research* 16: 239–262.

Crane M and Newman MC (2000) What level of effect is a no observed effect? *Environmental Toxicology and Chemistry* 19: 516–519.

Faustman EM and Omenn GS (1996) Risk assessment. In: Klaassen CD (ed.) *Casarett and Doull's Toxicology,* 5th edn., pp. 75–88. New York NY: McGraw-Hill.

Finney DJ (1978) *Statistical Method in Biological Assay*, 3rd edn. London: Charles Griffin and Company.

Gad SC and Weil CS (1988) *Statistics and Experimental Design for Toxicologists*. Caldwell, NJ: Telford Press.

Hamilton MA, Russo RC, and Thurston RV (1977) Trimmed Spearman–Karber method for estimating median lethal concentrations in toxicity bioassays. *Environmental Science and Technology* 11: 714–719.

Newman MC (1995) *Quantitative Methods in Aquatic Ecotoxicology*. Boca Raton, FL: CRC/Lewis Publishers.

Salsburg DS (1986) *Statistics for Toxicologists*. New York: Dekker.

Sprague JB (1969) Measurement of pollutant toxicity to fish. I. Bioassay methods for acute toxicity. *Water Research* 3: 793–821.

Sprague JB (1971) Measurement of pollutant toxicity to fish. III. Sublethal effects and 'safe' concentrations. *Water Research* 5: 245–266.

Suter GW, II (1993) *Ecological Risk Assessment*. Boca Raton, FL: CRC/Lewis Publishers.

US EPA (2002) *Methods for Measuring the Acute Toxicity of Effluents and Receiving Waters to Freshwater and Marine Organisms*, 5th edn. EPA-821-R-02e012. US Environmental Protection Agency, Washington, DC.

US EPA (2002) *Short-Term Methods for Estimating the Chronic Toxicity of Effluents and Receiving Waters to Freshwater Organisms*, 4th edn. EPA-821-R-02-013. US Environmental Protection Agency, Washington, DC.

Zhao Y and Newman MC (2003) Shortcomings of the laboratory derived LC50 for predicting mortality in field populations: Exposure duration and latent mortality. *Environmental Toxicology and Chemistry* 23: 2147–2153.