

Risk Assessment for Neurobehavioral Toxicity: SGOMSEC Joint Report

Dale Hattis,¹ John Glowa,² Hugh Tilson,³ and Beate Ulbrich⁴

¹Marsh Institute Center for Technology, Environment and Development (CENTED), Clark University, Worcester, Massachusetts; ²LMC/National Institute of Diabetes and Digestive and Kidney Diseases/National Institutes of Health, Bethesda, Maryland; ³U.S. Environmental Protection Agency, Research Triangle Park, North Carolina; ⁴Institut für Arzbeunuttek, Bundesgesundheitsamt, Berlin, Germany

Behavioral end points for neurotoxicity risk assessment have been developed and examined over the past three decades. They are now ready to move from simple qualitative guidelines, such as exemplified by reference doses, to more quantitative models, such as benchmark doses, based on dose-response information. Risk assessors, confronted by a wider array of methodologies and data than in the past, should be offered guidance in interpretation because now they have to deal with unaccustomed questions and problems. These include reversibility, susceptible populations, multiple end points, and the details of dose-response and dose-effect distributions. — *Environ Health Perspect* 104(Suppl 2):217–226 (1996)

Key words: benchmark dose, effect levels, identification of neurotoxicity, interindividual variability, quantitative models, reference doses, regulatory activity, reversibility

History

The demands for information on chemicals that might have adverse effects on the structure or function of the nervous system are changing. At the present time, there is

increased pressure for a more thorough screening for a wide variety of neurotoxic end points and an interpretation of the significance of effects on the nervous system, particularly in the context of premarket testing and quantification of neurotoxic risks that could contribute to setting overall priorities for risk reduction. This historical overview will explain how the professional community came to the current conception of the problem, how methodologies have evolved to meet the changing needs of decision makers, and how problem definitions and research techniques need to change further, both to meet future decision-making needs and to take advantage of opportunities to develop a fuller understanding of neurotoxic risks.

Risk assessment is the attempt to predict the likelihood that an agent will produce adverse effects in humans, typically from information obtained in animals. Assessment of risk dates from historical times, as it became the responsibility of authorities to protect its charges. However, it was not until recent times that this responsibility became formalized. The increase in the

number of chemicals to which people are exposed, the numbers of cases of adverse effects of those exposures, the different type of adverse effects, and the increased precision with which levels of compounds that may be related to those effects are assayed has placed an increased level of responsibility on the process of risk assessment.

Emergence of Neurobehavioral Toxicology

During the last 30 years, there has been a sharp increase in regulatory activity by various governmental agencies to protect the population from toxic agents. This regulatory emphasis is based on several factors, including the demonstrated causal relationships between adverse health outcomes and occupational/environmental exposures, realization of the pervasive presence of environmental exposures because of increased analytical chemical capabilities, and the concern that many chemicals in the environment have not been tested for potential toxicity. In the United States, at the federal level, four regulatory agencies have been given primary responsibility for regulating exposures to toxic chemicals: the U.S. Food and Drug Administration (FDA) (most food, drug, and cosmetic exposures), the Occupational Safety and Health Administration (OSHA) (workplace exposures); the U.S. Environmental Protection Agency (U.S. EPA) (air, water, pesticides, and miscellaneous exposures to industrial chemicals—the latter via the Toxic Substances Control Act), and the Consumer Product Safety Commission (CPSC) (1). Premarket screening is generally conducted for new food additives, drugs, pesticides, and some industrial chemicals when there is reason to suspect that anticipated use may pose unreasonable risks. In addition there is a provision for requiring testing of chemicals that are already on the market in the case of pesticides (on reregistration) and selected industrial chemicals. Similar requirements are present in a number of other countries.

Since 1975, several expert panels have recommended that regulatory agencies should screen chemicals for neurotoxicity; many of these recommendations specifically stated that neurobehavioral end points such as neurological functioning, motor activity and schedule-controlled behavior, be used in neurotoxicity hazard identification (2). Behavioral procedures are currently employed to evaluate

This joint report was developed at the Workshop on Risk Assessment Methodology for Neurobehavioral Toxicity convened by the Scientific Group on Methodologies for the Safety Evaluation of Chemicals (SGOMSEC) held 12–17 June 1994 in Rochester, New York. Manuscript received 1 February 1996; manuscript accepted 17 December 1995.

Address correspondence to Dr. Dale Hattis, Marsh Institute Center for Technology, Environment and Development (CENTED), Clark University, 950 Main Street, Worcester, MA 01610. Telephone: (508) 751-4603. Fax: (508) 751-4600.

Abbreviations used: FDA, Food and Drug Administration; OSHA, Occupational Safety and Health Administration; U.S.EPA, U.S. Environmental Protection Agency; CPSC, Consumer Product Safety Commission; IPCS, International Programme on Chemical Safety; ADI, allowable daily intake; NOAEL, no observed adverse effect level; IRIS2, Integrated Risk Information System; ALS, amyotrophic lateral sclerosis; RfD, reference dose; LOAEL, lowest observed adverse effect level; UFs, uncertainty factors; MF, modifying factor; AUC, Area under a concentration-time curve; BMD, benchmark dose; 5HT, serotonin; MDMA, methylenedioxymethamphetamine; 3SD, three standard deviations (abnormal effect); MPTP, 1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine.

developmental neurotoxicity of drugs in Japan and the European Union (3), while neurotoxicity testing protocols are being considered by the Organization of Economic Cooperation and Development (4). The International Programme on Chemical Safety (IPCS) has recommended that chemicals be evaluated for potential neurotoxicity using behavioral tests (5) and is currently sponsoring a collaborative study to validate the use of a neurobehavioral screening battery for the routine assessment of chemicals for neurotoxicity (6). Testing guidelines for several neurobehavioral end points, including some from developmental neurotoxicity, have been published by the U.S. EPA (7). Neurobehavioral measures are also used in neurotoxicity hazard identification by other regulatory agencies in the United States, including the FDA, OSHA and CPSC (8). Neurobehavioral end points are frequently used in conjunction with other measures of neurotoxicity such as neuropathology and neurochemistry in an integrated approach to assess chemicals for possible neurotoxicity (7).

The development of quantitative methods in neurobehavioral toxicity assessment clearly has its roots in behavioral pharmacology. Pavlov, Skinner, and Freud each contributed in their own way to the assessment of the behavioral effects of agents, and, as this type of analysis grew, behavioral pharmacologists quickly learned the value of complete dose–effect data from pharmacologists. While the era of quantitative analysis of behavioral pharmacology data was ushered in by the seminal work of Dews (9), quantitative work in behavioral toxicology closely followed. In one of the first demonstration of these methods Armstrong et al. (10) demonstrated behavioral changes resulting from exposure to mercury vapor using schedule-controlled behavior.

Development of More Quantitative Conceptions of Risk

Until recently, neurobehavioral end points have primarily been used in a qualitative sense to determine whether chemicals may produce neurotoxicity; i.e., hazard identification. During the last few years, there have been attempts to develop more quantitative descriptions of risk so that regulatory agencies might determine the level of exposure to a toxic chemical that should be considered safe in a regulatory context and the likelihood and possible extent of harm that might result from exposure to the chemical at different levels and durations.

The overwhelming majority of current attempts to quantify risk are based on concepts developed over 40 years ago in which a specific number, the allowable daily intake (ADI) is derived from experimental data to determine safe levels for regulatory purposes (11). As updated in present usage this approach is essentially based on finding a dose that does not produce an adverse effect in an experimental system (the no observed adverse effect level NOAEL) and that uses uncertainty factors to account for considerations such as differences between humans and animals, differences among humans in individual sensitivity, possible interactions between chemicals in the diet, and various departures from ideal quality and coverage of sensitive end points in the data base. The ADI is intended to represent the dose rate for a chemical (in mg/kg body weight/day) that is likely to be without appreciable risk of deleterious effects during a lifetime of exposure.

At the present time, risk assessment guidelines published by the CPSC (12), and under review by the U.S. EPA, discuss several techniques to obtain a quantitative estimate of risk. In spite of the lack of formal risk assessment guidelines, the U.S. EPA has used neurotoxic end points to make regulatory decisions for a substantial number of chemicals. For example, the EPA's Integrated Risk Information System (IRIS), which contains information on regulatory decisions made by the agency, indicates that approximately 20% of the nearly 550 standards or advisories in the database are based wholly or in part on neurotoxic end points. Approximately one half of the neurotoxic end points used were neurobehavioral effects, including neurological signs, changes in motor activity, and altered sensorimotor function (13).

In recent years there have been increased pressures for quantification of risk levels arising from *a*) the desire to evaluate the health benefits in some terms that can be compared with economic costs of regulatory control actions, *b*) priority ranking schemes to direct agency attention to arguably more significant problems (14,15) and *c*) needs for more quantitative assessments of significance in the context of ecological risk assessment.

Finally, advances in our scientific understanding now indicate that some neurotoxic effects in the human population may be subtle but significant. Such effects include

- Reversible impairments in sensorimotor function due to solvent exposures that

might contribute to industrial and automobile accidents,

- Cognitive/learning deficits resulting from developmental exposures to neurotoxins exposure to which cannot be entirely eliminated, but for which some relatively costly measures are possible for further reduction—lead, methyl mercury, PCBs.

- Contributions to delayed and cumulative neurodegenerative changes underlying Parkinsonism, and possibly Alzheimer's disease and amyotrophic lateral sclerosis (ALS).

Current Methodologies: Adequacies and Inadequacies

The Reference Dose Approach—The Default Position

Quantification of risk using the reference dose (RfD) approach involves estimation of the quantitative relationship between the magnitude of a response and the exposure or dose. In the RfD model, a high dose is usually one that produces some generalized or systemic effects, while a low dose usually produces no detectable response above background [i.e., it is a no effect level (NOEL), or a NOAEL if a judgment is made that a particular detectable effect is not "adverse"]. As described by Barnes and Dourson (16), the RfD is "an estimate of a daily exposure to the human population that is likely to be without appreciable risk of deleterious effects during a life-time." The RfD approach for noncancer end points such as neurotoxicity assumes that there is a threshold for the critical effect.

The RfD approach depends on the selection of a critical effect or the effect observed at the lowest dose level from a set of studies on the agent. Based on the available data, a NOAEL based on the critical effect is determined. If a NOAEL cannot be determined, then the lowest dose at which the critical effect is observed (the LOAEL) is determined. To obtain an RfD, the NOAEL is divided by a series of uncertainty factors reflecting various sources of uncertainty in the data set.

Table 1 lists guidelines for the use of uncertainty factors currently recommended by the U.S. EPA. Examples of uncertainty factors, which are usually factors of 10, include variation in sensitivity within human populations, extrapolation from animal data to humans, and less than lifetime to lifetime exposure. In the case where a LOAEL must be used because a NOAEL is unavailable, another factor of 10 is included. A modifying factor of less than 1- to 10-fold may be used

Table 1. Guidelines for the use of uncertainty factors in deriving reference dose (RfD).^a

Standard uncertainty factors		Guideline
H	Human to sensitive human	Use a 10-fold factor when extrapolating from valid experimental results from studies using prolonged exposure to average healthy humans. This factor is intended to account for the variation in sensitivity among the members of the human population.
A	Animal to human	Use an additional 10-fold factor when extrapolating from valid results of long-term studies on experimental animals when results of studies of human exposure are not available or are inadequate. This factor is intended to account for the uncertainty in extrapolating animal data to the case of average healthy humans.
S	Subchronic to chronic	Use up to an additional 10-fold factor when extrapolating from less than chronic results on experimental animals or humans when there are no useful long-term human data. This factor is intended to account for the uncertainty in extrapolating from less than chronic NOAELs to chronic NOAELs.
L	LOAEL to NOAEL	Use up to an additional 10-fold factor when deriving an RfD from a LOAEL instead of NOAEL. This factor is intended to account for the uncertainty is extrapolating from LOAELs to NOAELs.
Modifying factor		Use professional judgment to determine another uncertainty factor (MF) that is less than or equal to 10. The magnitude of the MF depends upon the professional assessment of scientific uncertainties of the study and data base not explicitly treated above, e.g., the number of animals tested. The default value for the MF is 1.

^aAdapted from Barnes and Dourson (15).

to account for the incompleteness of the data set (17).

The traditional RfD/ADI/Safety Factor approach has several advantages:

- It is relatively straightforward to apply and does not require complicated model building or analysis
- Through thousands of applications in the past, it is not yet known to have led to catastrophic adverse effects in humans.

Although the RfD/ADI/Safety Factor approach is widely used in quantitative risk assessment, the RfD approach has several potential limitations. By definition, the RfD approach uses only the NOAEL/LOAEL from the data set, and the NOAEL/LOAEL must be one of the doses used in the experiment. It is, therefore, possible that the number and spacing of doses in a study could affect the dose eventually selected as the NOAEL. As shown by Kimmel (17), a NOAEL could vary by a factor of 3 depending on the doses used in the experiment.

A potential problem with the RfD approach is that it does not take into account the shape of the dose response curve. Crump (18) described an experiment in which a steep dose-response curve resulted in a NOAEL well below that from an experiment having a dose-response curve with a shallower slope. Thus the RfD procedure could result in a more conservative estimate of risk for a chemical with a steeper slope, even though the steeper slope may indicate less interindividual variability in response.

The RfD approach also does not take into account the statistical uncertainties in determining the NOAEL from the experiment. For example, using a larger number of subjects per group increases the power of the experiment to detect an effect with statistical significance at a lower dose. Therefore, using larger groups of animals would tend to lead to a lower RfD. Similarly, measures having an inherently high baseline incident of a toxic effect could result in higher RfDs than measures that have a lower baseline effect incidents.

The RfD also makes the assumption that there is no increase in the incidence of toxicity over background at exposure levels at or below the RfD. It does not, therefore, provide for a procedure to determine risk at any other dose. This is particularly problematic if it is desired to assess *a*) the increase in risk that may be present if exposure occurs slightly above the RfD, or *b*) the confidence that a decision maker should place in the estimate that risk (incidence of effect) will be below specific numerical targets (or, conversely, the degree of uncertainty that might exist that risk might be above desired target levels).

There is at present no distinctive treatment of neurobehavioral effects in this process. Developmental effects are sometime given an extra factor although there is no general policy. For the long-term future development of neurobehavioral toxicology and decision making, the simple uncertainty

factor approach has a number of other disadvantages (19):

- No one knows how protective the RfD really is, either in general or in specific cases. What fraction of the diverse human population can be expected to experience adverse effects when exposed at the level calculated to be acceptable under the formula? In general, there may be some finite fraction of individuals who, because of disease or other causes, are marginal for biological functions affected by the chemical and who may be pushed beyond a functional threshold for an adverse effect by a small finite dose of the chemical. For example, for healthy workers there may indeed be a functional reserve capacity for oxygen delivery to the myocardium and hence a finite tolerance for a small impairment of oxygen-delivering capacity for the blood due to carbon monoxide. However, for a worker who has just begun to experience a myocardial infarction, oxygen delivery to portions of the myocardium is known to be seriously compromised, and it is possible that a small difference in oxygen delivery capacity due to a modest blood carboxyhemoglobin concentration could prove the difference between life and death for portions of the heart muscle that are suddenly forced to rely on collateral arterial vessels for oxygen supply

- The RfD procedure incorporates one specific social-policy standard for setting acceptable levels without making clear where the technical analysis leaves off and the policy/value analysis begins. Just how much risk of what degree of response in what proportion of people can we say is how unlikely? Is there no difference among different regulatory contexts and or different types of end points in the procedure that should be used to determine the level of exposure that is safe enough to permit?

- There is no defined or obvious way to incorporate newer types of relevant data on human interindividual differences in

- Rates of uptake/absorption for a constant environmental exposure (exposure variability)

- Rates of activating or detoxifying metabolism and excretion, producing differences in the concentration \times time of active metabolites per unit of absorbed dose at the site of toxic action (pharmacokinetic variability)

- Differential risk of response (response variability) for a given concentration \times time of active metabolites at the site of toxic action.

Table 2. Preliminary estimates of interindividual variability in various parameters likely to be related to susceptibility to different kinds of effects/chemicals.^a

Parameter and parameter class	Width of the 90% range ^b for an average chemical or test	Width of the 90% range ^b for chemicals/tests with greater interindividual variability than 95% of all others examined
Systemic pharmacokinetic parameters ^c		
Half-life for elimination (by metabolism and/or excretion)	2.3-fold	5.8-fold
Maximum blood concentration (after a single dose)	2.3-fold	11-fold
Area under a concentration–time curve (AUC)	3.0-fold	8.1-fold
Preliminary estimates of overall variability in sensitivity for neurotoxic risks		
Death from agents metabolized by plasma cholinesterase	5.5-fold ^d	Not available
Death from parathion (metabolized by paroxonase)	12-fold ^d	Not available
Effects of methylmercury poisoning in adults (6 end points)	12-fold ^e	78-fold ^e
Fetal/developmental effects of methylmercury poisoning (5 end points)	Approx. 460 ^b	Approx. 10,000 fold ^b

^aAdapted from Ashford et al. (20). ^bThe numbers are the ratios of the value of each parameter for the 95th percentile individual to the value of the parameter for the 5th percentile individual (see Figure 1 for illustration), based on the assumption that the values of the parameter are lognormally distributed (that is, the logarithms of the parameter values are normally distributed). ^cData from Hattis et al. (21). ^dData from Hattis et al. (22). ^eData from the National Academy of Sciences (19). ^fData from Hattis and Silver (23) and Hattis (24). Caveat: Imperfection of the dosimeter used here (maternal hair mercury) may have led to serious overestimation of the actual extent of interindividual variability.

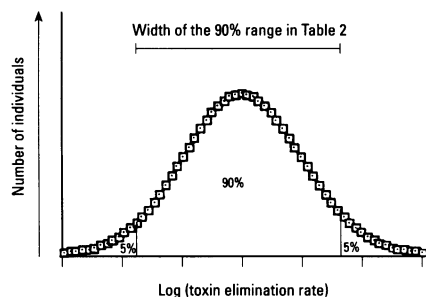


Figure 1. Interindividual variability of toxin elimination from the body based on the assumption that the logarithms of half-lives for elimination of toxins from the body are normally distributed.

In particular, it is also likely that the inability of the uncertainty factor paradigm (as usually formulated) to incorporate newer types of relevant information into a systematic procedure for updating assessments of health hazards has tended to discourage both the collection and analysis of potentially important data. One example of this is information on human interindividual variability in parameters that could affect susceptibility. Table 2 gives an overview of preliminary estimates of the extent of human interindividual variability in susceptibility for a number of parameters of potential interest for neurotoxicity risk

assessment. The numbers in both columns indicate the spread of interindividual variability observed, expressed as the ratio of a 95th percentile individual value to a 5th percentile individual value (Figure 1). The difference between the two columns can be thought of either as variability among chemicals tested in the same way (or, in some cases, different tests for the same chemicals), or the uncertainty facing an analyst in assessing the amount of interindividual variability in susceptibility to the carcinogenic effect of a chemical for which there are no direct measurements of human variability. The first column of numbers in Table 2 relates to median chemicals analyzed for each effect, or the median of several available tests of the same parameter, such as breathing rates; the second column represents observations for a 95th percentile high-variability chemical or test. Such data are not commonly compiled in this form in the normal course of risk evaluation for neurotoxic and other noncancer agents.

Alternative Approaches for Development of More Quantitative Information on Risks—Pros and Cons

Benchmark Dose/Crump. The approach originally proposed by Crump (18) uses the benchmark dose (BMD) in place of the NOAEL to determine the RfD. In that

sense, among the approaches surveyed here, the BMD makes the least disruption of current approaches for analysis of noncancer end points. The BMD is defined as the lower 95% (bound) confidence limit on the dose corresponding to a particular incidence of a quantal effect (usually 10 or 5%) in a dose–response curve calculated by fitting a mathematical model to the observed data. The dose–response models used for this purpose can be of any of a number of forms; those originally recommended by Crump include the polynomial forms often used for carcinogenesis risk assessment. As in the case of the RfD approach, uncertainty factors are applied to the defined benchmark dose to calculate the benchmark-RfD. It has been suggested that the BMD provides a common starting point for applying uncertainty factors and might result in RfDs that provide more comparable levels of protection than when NOAELs are used (17).

Advantages of the BMD include the fact that the approach uses data from the entire dose–response curve, and the statistical uncertainties arising from experimental design (dose spacing, numbers of animals used, background incidence of effects) all naturally affect the width of the confidence intervals around the points in the fitted dose–response relationship and therefore the degree of conservatism built into the risk assessment. Unlike the NOAEL approach, there is no requirement for the BMD to be one of the experimental doses. The BMD approach incorporates information on the shape of the dose–response curve and is not inconsistent with the previous assumption of a population threshold dose. More controversially, the BMD also allows for the estimation of risk at given levels of exposure. If exposure exceeds the RfD, an upper bound on the excess risk can be estimated, and some regulatory decisions can be made as to whether the potential change represents an unacceptable risk in a particular decision making context.

The BMD approach still utilizes uncertainty factors to calculate the RfD. Unlike the default RfD approach, however, which uses uncertainty factors of 10 to account for interspecies variability, less-than-lifetime exposure, and animal-to-human extrapolation (Table 1), some advocates of the BMD approach suggest that it should be used to define limits of acceptable risk (i.e., 1/10, 1/100, 1/1000, or 10,000 (25)). Thus, both the default RfD and BMD approaches rely on relatively arbitrary units of 10 for uncertainty factors that are used to calculate the RfD.

Gaylor/Slikker. Gaylor and Slikker (26) have proposed an approach to risk assessment that is based on the control variability of the measure of interest. In their preliminary presentation of the method, an example in which monkeys were treated with MDMA was used to demonstrate the risks of decreasing serotonin (5HT) levels. In that example, the normal variation in 5HT was determined from a control sample, and a level of 3 standard deviations (SD) was used as a metric to define an abnormal level of 5HT. The effects of several treatments with different dose levels of MDMA served to establish a dose-response function for the relationship between the decrease in 5HT as a function of MDMA treatment. On this dose-response function, a dose was chosen as one for which risks would be determined. The assumption that variation in levels of 5HT after this treatment was applied to the mean effect, obtained from the function, and the overlap of this distribution with the level at which 3SD in the control data provided an indication of the proportion of the population expected to be affected abnormally at that dose. The method is easily adapted to assessing the increase in risk posed by progressively smaller doses by comparing similar distributions of effect, until the additional area (over background) under these distribution curves beyond the 3 SD cutoff is as small as 1/100 or 1/1000.

The method has not been used extensively to fully evaluate potential concerns. One concern that has been voiced, however, is that the assumption that variability in effect exhibited at an interpolated dose should be measured rather than assumed. The use of assumed variability measures at very low doses may also preclude the ability to realistically predict effects at low doses. A clear advantage of the method is that it unambiguously defines an abnormal effect (3 SD), and as such may be useful in adapting risk assessment procedures to continuous variables that do not span a range that can be characterized by traditional scales.

Dews/Glowa. Dews (27) proposed a different method of assessing risks. In his approach, effects in individual animals are obtained such that a small effect can be defined in each subject. Typically, this was done by establishing a dose-effect function for several different subjects, using the simplest model that could be used to describe the function (typically a linear function was fitted). From each function a small effect, similar to that proposed by Crump (18) was chosen and

the mean and distribution of the effect was characterized.

Dews assumed that these individual differences in effect (or a log transformation) would be normally distributed. By making this assumption, it was easy to compare the distribution of expected effects using normal curve statistics (i.e. the probability of an effect in the population is distributed as a Z-score). By selecting a probability of interest (e.g., 1/10, 1/100, or 1/1000), the intercept of that point on the distribution with the abscissa specifies the dose at which that incidence of effect is expected.

The method has the advantage of being simple and relatively independent of the type of function chosen to represent the individual dose-effect functions. It also is one of the few methods that actually determines a population-sensitivity measure and uses it for risk assessment. The assumption that the doses producing specific amounts of effects are normally or lognormally distributed is an additional benefit, because a great deal is known about the normal curve, in contrast to the fact that we know very little about the shape of the dose-effect function at low levels. The level chosen is somewhat arbitrary, but 10% usually lies within the directly measurable portion of the curve.

A distinct disadvantage of the approach originally described by Dews (27) is that it may not be applicable to all types of data. For example, agents with irreversible effects, or developmental or chronic studies, rarely provide single subject data. Recently Bogdan et al. (unpublished data) described a method that remedies this deficit. In this approach, group design dose-effect data are obtained, i.e., a control group is dosed with vehicle and several dose groups are given a single dose each. From these data an iterative computer program generates all possible lines that can be fit to all possible combinations of one point per dose curve. Thus, a study with two animals per dose and three doses would generate eight possible combinations. The effects can be calculated in absolute terms or as percentage of the control group.

Probit and Logit Tolerance Distribution Models. When the available data on effects are quantal (plus/minus; cases/non-cases) and the primary issue for analysis is the spread of interindividual variability in a population, log probit and logit dose-response functions may be appropriate. Both the log probit and the logit are tolerance distribution models: they are based on the assumption that the effect in question is produced

in different individuals when individual threshold doses are exceeded. In the case of the log probit model, these threshold doses are assumed to be lognormally distributed in the exposed population (that is, the logarithms of the threshold doses for individual people/animals are assumed to have a normal gaussian distribution). In the case of the logit model, the population distribution of thresholds is assumed to correspond to a logistic function.

In practice there is very little difference in the dose response projections made using these two models for effect incidences between 1 and 99%. The logistic model has slightly broader tails and therefore makes somewhat larger predictions of low dose risks when projections are made from high-dose observations. A report by the U.S. National Academy of Sciences (19) has a comparison of low dose predictions of risks from fetal developmental exposure to methylmercury using the logit (28) and log probit models, as applied to the human data from the mass poisoning episode in Iraq (29). At a dose predicted by the logit model to produce a 0.5% incidence of late walking in gestationally exposed offspring, the probit model prediction was for a 0.25% incidence of the effect.

The logit model tends to be favored by statistically oriented workers because of its computational convenience in fitting data. The traditional maximum likelihood procedure of Finney (30) often used to fit the log probit function, is by contrast relatively cumbersome, although automated routines are available. The rationales advanced by advocates of the probit model include *a*) at least a vague mechanistic foundation (a log-normal distribution of thresholds would be expected to be produced if there were several factors that contributed to interindividual differences in sensitivity and if these factors tend to act multiplicatively in determining individual thresholds); and *b*) a long history of prior usage in the traditional analysis of animal dose-response data (most notably in the calculation of LD₅₀ values). For example, Gaddum (31) described one of the first quantitative approaches to risk assessment in his early attempt to develop a method to establish safe doses. While a number of parameters of the dose-effect data were used, the approach was essentially to establish a safe dose as one calculated to be six standard deviations below an easily observable effect (e.g., LD₅₀). Of note, it set the occasion for displacement in the dose dimension based on variability established in the effect dimension and, more importantly,

used a fair portion of the available dose–effect information. However, it was a point estimate that provided no indication of this uncertainty in the risk (safe) figure.

Although both of these models are based on an assumption that effects are produced in different individuals when their personal thresholds are exceeded, in contrast to the RfD procedure, there is no necessary assumption of a population threshold—a dose so low as to be below the threshold of every single person in a large mixed population. However, particularly for the probit model, expected effect incidences decline to very low levels after doses are reduced below several SDs from the population median threshold.

Finally, as implied at the start of this subsection, the assumptions underlying the probit and logit models are rarely compatible with the mechanisms that produce the kind of continuous data common in describing many neurotoxicological end points. An exceptional case of the application of the probit model to continuous (nonquantal) data on the depression of dopamine concentrations in specific brain regions by MPTP is based on the assumption that the measured reductions in dopamine levels may reflect the amount of killing of relevant cells by MPTP (24). Such killing could be expected to result from the exceedance of individual thresholds, where the individuals in this case are individual cells in a mixed population of neurons.

The basic idea that at some level there should often be thresholds for individual responses for many types of adverse effects does have some mechanistic justification. The dominant paradigm of traditional toxicology and pharmacology views biological systems as complex interacting webs of processes so designed that the perturbation of any one parameter automatically gives rise to countervailing influences that tend to keep the system within normal limits (32).

Given that some effects are produced by individual threshold processes, the probit model offers a direct opportunity to assess the spread of the population distribution of the thresholds. The estimated SD of the population distribution of the logarithms of the thresholds is the reciprocal of the probit slope in conventional plots of the probit of response versus log dose (24). A caveat here is that, when applied to animal data, this SD refers to the animal population. It is likely that relatively genetically uniform strains of animals, raised under controlled laboratory conditions, may often have less interindividual variability than outbred

humans exposed at different ages and with different histories of other exposures, nutrition, etc. Because of this, there may be a need to adjust expectations of interindividual variability in humans when using animal data to project human risks with tolerance distribution models.

Quantitative Mechanistic Modeling.

The interplay between mathematical theory and experimental neurotoxicology could be greatly enriched through efforts to build more mechanistic understanding into the mathematical models used to represent neurobehavioral changes. A paper elsewhere in this volume (24) emphasizes the following:

a) The mathematical form for a dose–time–response model is ideally not just a convenience for summarizing or fitting a particular data set—it represents a hypothesis. The more this hypothesis reflects a mechanistically sophisticated view of the likely reality, the more it can lead to potentially informative validating/invalidating types of predictions about the results of real experiments, and (in the long run) reasonably credible predictions outside the range of potential direct observation . . .

b) Models are simplifications of reality. They are intended to represent the likely behavior of a complex system of interest by focusing on the behavior of a few salient features or components of the system. If the features included in the model are the prime determinants of the behavior of the system (that is, if other variables that could also affect the system are relatively constant or relatively unimportant) then there is a hope that the model will reasonably accurately represent the behavior of the system over a particular domain of conditions.

Some features of neurotoxic effects that are amenable to quantitative dynamic modeling at the present time are

- The ability to establish definitive relationships between external dose and the internal dose–time profile at the presumed site(s) of action for some chemicals. This is best done using the well-developed paradigm of physiologically based pharmacokinetic modeling (33,34), which has been extensively used in carcinogenesis risk assessment in recent years (35). The importance of good active site dosimetry is to separate nonlinearities of pharmacokinetic origin from nonlinearities arising from fundamental neurotoxic processes in the chain of events leading from external exposure to the production of neurobehavioral responses

- The familiar tools of Michaelis-Menten enzyme kinetics are readily adaptable to elucidating the dose response

relationships for both saturable transport processes and the inhibition of enzymes that are important for the survival and functioning of cells in the nervous system (36).

- The accumulation and repair of reversible neurotoxic damage can be presented in pharmacodynamic models in some cases (37).

Reference (24) includes some more exploratory uses of mechanistic models for the central processes involved in neurotoxicity including *a)* chronic accumulation of irreversible damage via the loss of neuronal cells; *b)* the use of intermediate biomarkers along the causal pathway to neural dysfunction to aid in dose–response and population effects modeling (38); and *c)* the implications of the structural features of neuronal systems (e.g., redundancy, series arrangement of cells performing functions, plasticity) for the relationships between dose–response relationships for the inhibition of neurobehavioral function and dose response relationships for the inhibition of individual neurons.

Recommendations for Assessing New or Old Agents with Good Data

The previous sections have described the generic advantages and disadvantages of several quantitative risk assessment methodologies. The current section will attempt to apply these models to various types of data sets that may be encountered in the risk assessment process. In this section it is assumed that some hazard identification has already occurred and that the risk assessor is presented with a set of data requiring a risk assessment decision. A preliminary assessment of the data set has indicated that a neurobehavioral effect is the critical end point and quantitative risk assessment should proceed using these data.

Typically the first step in the process is to identify the relevant toxicity studies and the supporting data that can be used to select and modify the best approach for the generation or application of a dose–effect function. The most usable types of dose–effect functions are those that clearly delineate control variability and assess a sufficient number of doses to *a)* establish a NOEL/NOAEL or a benchmark dose, *b)* characterize the slope of the dose–effect function; and, ideally, *c)* relate to potentially testable theories of toxic mechanism. The application of risk assessment methodologies clearly depends on the type and amount of data available. The simple demonstration

of an effect may not be sufficient to assess risks quantitatively.

Examples

As with the RfD approach, several features of the type of end point used for assessment, the nature of the control data, the experimental design (e.g., number of doses, times), variability in effect, and mechanistic information can contribute to the risk assessment process using quantitative methods. In an effort to present the role of these variables and the possible use of the different quantitative methods under different conditions, the following representative examples of datasets are presented:

a) A developmental study was conducted in which groups of pregnant rats were treated with a penicillin derivative (100, 500, and 2000 mg/kg, po) and a control group was treated with vehicle during days 6 to 15 of pregnancy. A typical battery of developmental end points was assessed, and changes in these variables were used as a covariate in the determination of significant neurobehavioral effects. In this study a significant change in activity was found at the highest dose as determined using concurrent and historical control levels of activity. The dose had no toxic effects on the mother. Additional information came from a placental transfer study that established embryo-fetal exposure over the whole dose-range studied and from a pharmacokinetic study in pregnant animals that showed a linear increase in plasma levels up to the highest dose.

This developmental study provided good dose-effect data and a range of normal control data, allowing each of the methods to be applied. An RfD could be calculated based on the next-to-highest dose being chosen as a NOAEL, as it did not produce observable effects. This would, under the current example, produce an RfD that was considerably below therapeutic levels, indicating that at therapeutic levels some risk was present. The calculation of a risk figure by each of the other methods would also likely point to a dose of penicillin that might be expected to exhibit developmental risk at therapeutic levels. Use of the probit model for this purpose would be marginal, however, since it would require an assumption about the probit slope, and there is only one positive dose.

Risk assessment for medicinal products rarely involves the definition of a dose that is regarded as virtually safe. The ultimate requirement is that the minimal therapeutic dose (concentration of active substance

in most cases) doses not pose an unreasonable risk to the patients in relation to the anticipated therapeutic benefits. Preclinical studies, although generally referred to as safety testing, will be directed, therefore, at detecting and characterizing adverse effects with the aim of enabling the investigators to assess/quantify risks at therapeutic dose levels. During the preclinical phase of pharmaceutical testing the key elements of risk assessment for neurobehavioral/neurotoxic end points involve judging the relevance of findings to humans, considering the biological properties of the animal models used.

b) In another study, a collection of case reports in which an increased incidence of neurological symptoms associated with eosinia myalgia were noted in several countries where a tainted tryptophan product had been sold. Because this data set involved simply the observation that an increased incidence of effect was seen in an exposed population and neither dose-effect nor control data were available, none of the methods are applicable; more data would be required.

c) A case-control epidemiological study involved subjects with neurotoxic signs suspected to have resulted from deltamethrin (pyrethroid insecticide) exposure. Subjects matched for age, living standards, and other health and socioeconomic factors were used as controls. Effects were determined by clinical assessment, exposure was determined by questionnaire.

Because individuals in the control group in case-control studies are potentially exposed (yet unaffected) and it may be difficult to establish dose-effects in the exposed group, this is a potentially difficult design to use for risk assessment. However, if a stratified analysis reveals that effects increased over a range of exposures it may be possible to apply the benchmark methods or other dose-time-response models, ideally with corrections for the biasing effects of inaccuracies in the estimation of exposures (23).

d) In an occupational workplace study blood levels of toluene were determined directly after testing using an automated neurotoxicology test battery. Populations from several sites at the factory were tested, providing several ranges of level and including a nonexposed control. The performance deficits on the battery were positively correlated with hippuric acid level.

This workplace study would provide good effect and dose data, good characterization of human pharmacokinetics, and at

least some preliminary hypothesis generation for neurotoxic mechanisms, allowing each of the methods to be used.

e) A dose-response study in rats was designed to characterize acute effects of alprazolam on activity. Rats were placed in activity monitors once daily until stable baselines were established. Animals were given vehicle and different doses on separate widely spaced occasions until individual dose-effect curves were established for 10 animals. Control variability was approximately 10%. Activity was unaffected at the lowest dose tested and abolished at the highest dose tested.

This dose-response study would provide good effect and dose data, allowing each of the methods to be used. Since individual animal variability is directly assessed in this study, the method of Dews provides the most direct measures of risk. Probit and mechanistic analyses may also be possible.

f) Another dose-response study in rats was designed to characterize acute effects of MPTP on activity. Four groups of six rats each were treated with vehicle or with one of three doses of MPTP and tested three days later. Levels of activity were 200 photocell counts/15 min session in the control group, and they progressively decreased in a dose-related manner in each of the experimental groups. Variability was essentially the same in each of the groups.

This type of study would provide good effect and dose data, allowing each of the methods to be used. Since variability in effect was comparable between control and dosed groups, the method of Gaylor/Slikker could be applied without reservation. Because some aspects of the mechanism of action of MPTP have been characterized (e.g., loss of neurons relevant to some functions) a quantitative mechanistic analysis may also be possible.

g) Finally one study assesses the effects of chronic (90-day) exposure to *n*-hexane (0, 100, 300, 1000 ppm) using a functional observational battery (FOB) in mice. The battery was assessed twice—on the second day and the second week after the termination of exposure. Effects were noted on 10 end points 2 days after the termination of exposure in the high-dose group. These effects were not apparent 2 weeks after exposure ended.

This study would provide an initial indication of dose-effect that dissipated with retesting. Quantitative mechanistic analysis may be possible if sufficient data are available on the dynamics and dose-response of

Table 3. Applicability of risk assessment methods to the exemplary data sets.

Example Method	<i>a</i>	<i>b</i>	<i>c</i>	<i>d</i>	<i>e</i>	<i>f</i>	<i>g</i>
RfD	+	+	–	+	+	+	+
BMD	+	–	+	+	+	+	+
Dews	+	–	–	+	+	+	+
Gaylor	+	–	–	+	+	+	+
Probit/logit	–	–	+?	+	+	+	+
Causal/mechanistic	–	–	–	+	+?	+	+

+ Denotes applicable. – Denotes not applicable. ? Denotes possibly applicable but questionable.

either the appearance or the disappearance of the functional changes.

It is obvious that each of these quantitative risk assessment methods can be used under various conditions with different types of data and, also, some methods are inappropriate with some types of data. The applicability of the different models to the different data sets as described is summarized in Table 3. When faced with these types of choices, the risk assessor may be compelled to choose a method with which he/she is most familiar or one that produces the most conservative risk figure. Since few studies have directly compared these different methods using the same data set, such choices may be less than ideal. Rather, it may be more instructive to compare the figures produced by each to understand the methods better.

With the possible exception of the probit and mechanistic models, virtually none of the methods conceived after the development of the RfD actively contribute to the exploration of the effectiveness or necessity of uncertainty factors. Crump (18) and Gaylor and Slikker (26) explicitly state that uncertainty factors have to be added to the number produced by their methods. Dews (27) discusses the possibility of finding species more and less susceptible than humans and interpolating rather than extrapolating for interspecies differences. Many of these methods, especially that of Dews (27), explicitly measure intraindividual variability, which may discount the need for including that factor. Several other studies have discussed various assumptions related to inter-species uncertainty factors [e.g., Rees and Glowa (39); Rees and Hattis (32)] concluding that these factors can be measured rather than assumed. Acute-to-chronic uncertainty factors may also apply for some agents and not others—no one has ever accomplished a full-term lifetime study to approach this question directly. Thus, for each of the examples provided below, one or more uncertainty factors could be applied. The

risk assessor must recognize that the application of these factors may have considerably more weight on the numerical outcome than the differences between these methods discussed.

Emerging Issues

There are a number of other issues important for neurobehavioral toxicity that should be considered by the risk assessor in evaluating datasets.

Reversible and Irreversible Changes.

One issue of concern to the risk assessor is the observation that a toxicant-induced change in behavior might dissipate soon after cessation of dosing or appear to lessen during the course of repeated exposure. Persistent changes in neurobehavioral measures should be always be viewed with a high degree of concern, but it does not always follow that reversible changes should be discounted or ignored in the risk assessment process. It is important to realize that the nervous system has a combination of special features not found in other organ systems. For example, the nervous system is composed of a number of different types of cells, each having its own functions and vulnerabilities. After a certain age, the process of neurogenesis ceases, and toxic damage to the brain or spinal cord could result in permanent loss of nerve cells. If the loss is concentrated in one of the nervous system's functional subsystems, the outcome could be devastating. One of the key features of the nervous system is that it has the capability to compensate for loss of function following damage and this compensation could mask the presence of nervous system damage. In the peripheral nervous system, if the cell bodies are not damaged, the axons have the ability to regenerate and attempt to reach their original target site. Therefore, eventual return of sensation and function might gradually occur after toxicant exposure. Residual damage to the peripheral nerve might still be present, however, and detected only by using tests capable of detecting relatively subtle neurobehavioral

deficits. Neurons in the central nervous system also have the ability to regenerate, but they have a much more difficult task in reaching their original targets due to both the presence of scar tissue formed by proliferating glia and to the increased complexity of the connections in the central nervous system. Loss of neurons in the central nervous system is generally regarded as permanent.

At the present time, there is limited understanding about how to calculate or correct for compensation or reversibility in neurotoxicity risk assessment. Biologically based dose response models are being developed to assist the risk assessor in evaluating such changes following exposure to neurotoxic agents (40).

Multiple End Points in Neurobehavioral Studies. There are several types of neurobehavioral effects that can be measured following exposure to toxic agents (8,41). In the case of humans, there are a number of examiner-administered and paper and pencil tests used to assess sensory, motor, cognitive, affective, and personality states or traits. Human neurobehavioral toxicology has also adopted a number of techniques from neurology and neuropsychology to assess nervous system impairment. Likewise, behavioral procedures from experimental psychology, behavioral pharmacology, and neurology are often used to detect and characterize neurobehavioral toxicity in animals. Frequently, neurobehavioral assessments use a number of measures in a battery of tests. Such batteries can consist of different kinds of data including continuous, categorical, and rank data. The risk assessor must be concerned about the different levels of power inherent in each of the different types of test measures since that could significantly affect the end point selected as the critical effect in the quantitative risk assessment. Furthermore, the risk assessor should be aware that the use of multiple end points in the same animals or subjects might need to be corrected statistically to avoid false positives.

Susceptible Populations. In general, we are all at risk of being adversely affected by exposure to neurotoxic agents. Individuals of certain age groups, health states, and occupations, however, may be at a greater level of risk. It is widely accepted that the developing nervous system is differentially sensitive to chemical insult. During the developing period, the nervous system is actively growing and establishing the necessary connectivity for normal functioning during adulthood. Protective systems such as the blood-brain barrier and detoxification

mechanisms such as certain metabolic enzymes may not be present during critical periods of development. Therefore, exposure to chemicals during development could result in a range of effects that would not be present or detectable in the adult organism.

It is now evident that the presence of developmental neurotoxicity may depend on the specific periods of nervous system development (42). In addition, the results of early developmental exposure may become evident only as the nervous system matures and ages. With aging, the level of risk for a number of health-related factors increases and it is possible that the risk for toxic insult to the nervous system may also increase. It has been hypothesized that the aging nervous system may have a decreased ability to compensate for toxic insults and that exposure to neurotoxicity could increase the rate of age-related neuronal cell loss (43). Examples of other possible susceptible populations include individuals having neurodegenerative disorders such as Parkinson's or Alzheimer's Disease, those with nutritional deficiencies and certain ethnic groups. Gender-dependent responses to neurotoxicants have also been noted.

Most current standard risk-assessment methodologies attempt to correct for differences in susceptible populations by using an uncertainty factor for within population variability. Whether a factor of 10 or less is sufficient to protect against the

differential sensitivity of different groups is not known, and the risk assessor should be aware that qualitative, as well as quantitative, differences could affect the risk assessment process.

Shape of the Dose-Response Curve.

Many of the methods used in quantitative risk assessment assume a linear or sigmoidal monotonically increasing dose-response curve. However, inverted U-shaped dose-response curves are sometimes observed in neurobehavioral studies. For example, it is sometimes observed that an agent will increase the frequency of a response, such as motor activity or schedule-controlled behavior, at low doses while decreasing responses at higher doses. In such cases, it may be difficult to determine which part of the dose-response curve should be used in quantitative risk assessment. The risk assessor should be aware that different critical effects could be obtained depending on the arm of the dose-response curve which is selected and that some quantitative methodologies will not be appropriate for such data.

During the process of risk recognition and from attempts to test for risk routinely to prevent or regulate exposure to neurotoxicants, we appear to have discovered valuable information about the behavior of investigators and risk assessors. In brief, there appears to be less interaction than would be desirable.

Summary

The concept of risk as it relates to neurobehavioral toxicology has evolved over time. Guidelines for how to test chemicals for neurotoxicity are being followed by risk assessment guidelines. The focus in research has shifted from qualitative risk assessment, i.e., hazard identification, to quantitative risk assessment, i.e., how to predict risk based on available data.

Several quantitative models of risk assessment have been proposed during the last several years. The traditional default approach, i.e., the RfD, is based on the allowable daily intake hypothesis proposed over 40 years ago. Risk assessors are now confronted with a wide array of risk assessment methodologies, each having their own advantages and disadvantages, depending upon the experimental design and characteristics of the dataset.

Current risk assessments using neurobehavioral data face several problems that could confound the interpretations of the result, including reversibility of effect, susceptible populations, multiple end points, and the shape of the dose-effect curve. Clearly, the state-of-the science for quantitative risk assessment of neurotoxicity is evolving and can be improved only through additional research.

REFERENCES

1. Fisher F. Neurotoxicology and government regulation of chemicals in the United States. In: *Experimental and Clinical Neurotoxicology* (Spencer P, Schaumburg H, eds). Baltimore: Williams and Wilkins, 1980;874-882.
2. Tilson H. Neurotoxicology in the 1990s. *Neurotoxicol Teratol* 12:293-300 (1990).
3. Kimmel RA. Current status of behavioral teratology: science and regulation. *Crit Rev Toxicol* 19:1-10 (1988).
4. Organization of Economic Cooperation and Development. Summary report OECD Ad Hoc Meeting on Neurotoxicity Testing. of Eastern Research Group, Lexington, MA, 1990.
5. WHO. Environmental Health Criteria 60: Principles and Methods for the Assessment of Neurotoxicity Associated with Exposure to Chemicals. Geneva:World Health Organization, 1986.
6. Moser VC, MacPhail RC. International validation of a neurobehavioral screening battery: the IPCS/WHO collaborative study. *Toxicol Lett* 64/65:217-223 (1992).
7. U.S. Environmental Protection Agency. OPPTS Test Guidelines: Neurotoxicity. Springfield, VA:National Technical Information Service, No PB91-154617, 1991.
8. U.S. Environmental Protection Agency. Final report: Principles of neurotoxicity risk assessment. *Fed Reg* 59:42360-42404 (1994).
9. Dews PB. Modification by Drugs of Performance on Simple Schedules of Positive Reinforcement. *Ann NY Acad Sci* 65:268-281 (1956).
10. Armstrong RD, Leach LJ, Belluscio PR, Maynard EA, Hodge HC. Behavioral changes in the pigeon following inhalation of mercury vapor. *Am Ind Hyg Assoc J* 24:366-75 (1963).
11. Lehman AJ, Fitzhugh OG. 100-Fold margin of safety. *U.S. Quart Bull* 18:33-35 (1954).
12. Consumer Product Safety Commission. Labelling requirements for art materials presenting chronic hazards; guidelines for determining chronic toxicity of products subject to the FHSA; Supplementary definition of 'toxic' under the Federal Hazardous Substances Act; final rules. *Fed Reg* 57:46626-46674 (1992).
13. Tilson HA, MacPhail RC, Crofton KM. Setting exposure standards: A decision process. *Environ Health Perspect* 104(Suppl 2):401-405 (1996).
14. U.S. EPA. *Unfinished Business: A Comparative Assessment of Environmental Problems*. Washington:U.S. Environmental Protection Agency, 1987.
15. U.S. Environmental Protection Agency. *Reducing Risk: Setting Priorities and strategies for Environmental Protection*. Washington: U.S. Environmental Protection Agency, 1990.
16. Barnes DG, Dourson M. Reference dose (RfD): description and use in health risk assessments. *Regul Toxicol Pharmacol* 8:421-486 (1988).
17. Kimmel CA. Quantitative approaches to human risk assessment for noncancer health effects. *Neurotoxicology* 11:189-198 (1990).

18. Crump KS. A new method for determining allowable daily intakes. *Fundam Appl Toxicol* 4:854–871 (1984).
19. National Academy of Sciences. *Seafood Safety* Washington:National Academy Press, 1991.
20. Ashford NA, Bregman C, Hattis DB, Karmali A, Schabacker C, Schierow LJ, Whitbeck C. *Monitoring the Community for Exposure and Disease: Scientific, Legal, and Ethical Considerations*. Atlanta:Agency for Toxic Substances and Disease Registry, 1991.
21. Hattis D, Erdreich L, Ballew M. Human variability in susceptibility to toxic chemicals — a preliminary analysis of pharmacokinetic data from normal volunteers. *Risk Anal* 7:415–426 (1987).
22. Hattis D, Bird S, Erdreich L. *Human variability in susceptibility to anticholinesterase agents*. CTPID 87-4. Boston:Massachusetts Institute of Technology, December, 1987.
23. Hattis D, Silver K. Human interindividual variability—a major source of uncertainty in assessing risks for non-cancer health effects. *Risk Anal* 14:421–431 (1994).
24. Hattis D. The challenge of mechanism-based modeling in risk assessment for neurobehavioral end points. *Environ Health Perspect* 104(Suppl 2):381–390 (1996).
25. Kimmel CA, Gaylor DW. Issues in qualitative and quantitative risk analysis for developmental toxicology. *Risk Anal* 8:15–20 (1988).
26. Gaylor D.W. and Slikker, W. Risk assessment for neurotoxic effects. *Neurotoxicology* 11:211–218 (1990).
27. Dews PB. Estimation of low risks *The Pharmacologist* 22:159 (1980).
28. Cox C, Clarkson TW, Marsh DO, Amin-Zaki L, Tikriti S, Myers GG. Dose-response analysis of infants prenatally exposed to methylmercury. An application of a single compartment model to single-strand hair analysis. *Environ Res* 49:318–332 (1989).
29. Marsh DO, Clarkson TW, Cox C, Myers GJ, Amin-Zaki L, Al-Tikriti S. Fetal methylmercury poisoning. Relationship between concentration in single strands of maternal hair and child effects. *Arch Neurol* 44: 1017–1022 (1987).
30. Finney DJ. *Probit Analysis*. Cambridge, UK:Cambridge University Press, 1971.
31. Gaddum JH. The estimate of the safe dose. *Br J Pharmacol*. 11, 156–160 (1956).
32. Rees DC, Hattis D. Developing quantitative strategies for animal to human extrapolation. In: *Principles and Methods of Toxicology*, 3rd ed (AW Hayes ed). New York:Raven Press;1994; 275– 315.
33. Fiserova-Bergerova V. *Modeling of Inhalation Exposure to Vapors: Uptake, Distribution, and Elimination*, Vols 1, 2. Boca Raton, FL:CRC Press, 1983;162 pp, 176 pp.
34. Ramsey JC, Andersen ME. A physiologically based description of the inhalation pharmacokinetics of styrene in rats and humans. *Toxicol Appl Pharmacol* 73:159–175 (1984).
35. Hattis D. Use of biological markers and pharmacokinetics in human health risk assessment. *Environ Health Perspect* 89:230–238 (1991).
36. Hattis D. Pharmacokinetic principles for dose rate extrapolation of carcinogenic risk from genetically active agents. *Risk Anal* 10:303–316 (1990).
37. Hattis D, Shapiro K. Analysis of dose/time/response relationships for chronic toxic effects—the case of acrylamide. *Neurotoxicology* 11:219–236 (1990).
38. *Molecular Epidemiology: Principles and Practices* (Schulte P, Perera R. eds). New York:Academic Press, 1993.
39. Rees RC, Glowa JR. Extrapolations to humans for neurotoxins. In: *The Vulnerable Brain and Environmental Risks*. Vol 3: *Toxins in Air and Water* (Isaacson R, Jensen K Eds). Plenum Press, New York:1994; 207–230.
40. Hattis D, Crofton K. Use of biological indicators of causal mechanisms in the quantitative assessment of neurotoxic risks. In: *Handbook of Neurotoxicology*. Vol. 3 *Approaches and Methodologies* (Chang L, Slikker W, eds). New York:Dekker, in press.
41. Anger WK. Neurochemical testing of chemicals: impact of recommended standards. *Neurobehav Toxicol Teratol* 6:147–153 (1984).
42. Rodier PM. Time of exposure and time of testing in developmental neurotoxicology. *Neurotoxicology* 7: 69–76, (1986).
43. Weiss B. Risk assessment: the insidious nature of neurotoxicity and the aging brain. *Neurotoxicology* 11:305–313, (1990).