Colloquium on Approaches to Quantifying Health Risks for Threshold or Nonlinear Effects at Low Dose

September 28, 2000

Final Summary Report

Submitted to:
Risk Assessment Forum
U.S. Environmental Protection Agency
Washington, DC

Submitted by:
The CDM Group, Inc.
Chevy Chase, MD

Contents

Contents (continued)

Introduction

The Colloquium on Approaches to Quantifying Health Risks for Threshold or Nonlinear Effects at Low Dose took place September 28, 2000, at the Omni Shoreham Hotel in Washington, DC. The meeting was held to explore approaches to characterizing variability and uncertainty in RfDs/RfCs and to provide a probabilistic framework for estimating risks associated with exposures above the RfD in order to assist Environmental Protection Agency (EPA) economists in valuing health benefits associated with environmental regulations.

Outside experts made presentations on distributional analysis of threshold effect based on dose response; characterizing risks above the reference dose; expected values of population dose-response relationships; risk-based reference doses; use of categorical regression to characterize risk above the RfD; and risks between the LOAEL and the RfD. A discussion period followed the presentations, focusing on a number of questions submitted by participants.

Vanessa Vu, National Center for Environmental Assessment (NCEA) Assistant Director, and Al McGartland, National Center for Environmental Economics (NCEE) Director, served as colloquium chairs. External presenters are listed in Appendix A, and colloquium participants are listed in Appendix B. The colloquium agenda is provided as Appendix C.

Welcome

Bill Wood, Executive Director of EPA's Risk Assessment Forum, welcomed everyone to the colloquium. He noted that NCEA had asked the Forum to organize the colloquium to explore issues surrounding quantification of risk above the RfD. In addition, he said the purpose was different from similar events in the past that focused simply on risk assessment issues. The colloquium resulted from a partnership between NCEA and NCEE and included risk assessors, economists, and outside experts from academia, consulting firms, State environmental programs, and others. He asked participants to submit one or two questions to help focus the afternoon's discussion.

Background on Risk Assessment

Vanessa Vu, NCEA Assistant Director, provided background on the need for the colloquium, focusing on how risk assessment methods can be improved to enable better benefit valuation and analysis. Her presentation slides appear at the end of this report as Appendix D. She explained the structure of the meeting. The objective was to explore possible approaches for quantifying risk below the point of departure, especially for those substances with biological thresholds or nonlinear dose-response curves at low dose.

Dr. Vu noted that EPA has many efforts underway to improve risk assessment methods for mixtures, aggregate risk of single contaminants, and cumulative risk of multiple stressors. EPA is also trying to harmonize the approach to all human health endpoints by moving away from the

terms "cancer/noncancer effects." To do so requires improving the understanding of mode of action and its incorporation into risk assessment.

For most (although not all) of the chemicals EPA is concerned with, it is necessary to take available human data or toxicology data from animal studies and extrapolate to exposures of interest. The Agency is also developing approaches to consider different exposure scenarios, including chronic, acute, and episodic exposure, as well as duration and temporal response, and to better characterize risk among sensitive subpopulations.

EPA currently has two approaches for characterizing risk. In the first, probabilistic estimates are used with carcinogens. The second approach, described in the revised draft cancer guidelines, uses margin of exposure for carcinogens with a nonlinear dose-response curve. The Agency assumes there is a biological threshold for reference dose, reference concentration, and margin of exposure for noncancerous health effects.

Dr. Vu showed a graph depicting a hypothetical dose response for a noncancer health endpoint. She explained that the Agency has used the 95 percent lower confidence limit to derive a benchmark for extrapolation to low dose because of considerations of statistical variability of the animals studied as well as the design of studies.

For effects believed to operate through a biological threshold, without knowing the dose-response curve from the observed range, the Agency's typical procedure is to apply uncertainty factors to account for extrapolation from animal data and across human populations in order to determine a reference dose or reference concentration. The day's speakers were to talk about this issue.

Dr. Vu noted that some EPA offices, such as the Office of Pesticides, use a margin of exposure approach for all noncancer effects. That approach takes any point of departure, such as an LED10 or one based on a NOAEL or LOAEL, accounts for potential estimates of human exposure, and determines the adequacy of the margin of exposure.

She showed another graph depicting a typical cancer dose-response curve based on observed data. When there is not enough information to know the mode of action or the shape of the dose-response curve, the Agency uses a linear default. If there is adequate mode of action information to suggest the shape of the dose-response curve but not to prove it, the Agency uses the margin of exposure approach, or an RfD-like approach when there are data to demonstrate a biological threshold mechanism.

Just as with the reference dose, a margin of exposure approach can be used when the anticipated human exposure level is known and can be compared with the LED10 to determine the adequacy of the margin of exposure.

Dr. Vu noted that while improving risk assessment methods is important as a way to improve risk characterization, it is also essential to providing health benefit analysts with adequate information to characterize benefits. She described the information economists need from risk assessors in order to perform cost-benefit analysis: a full characterization of the range of health effects potentially associated with contaminants, including the nature of a specific effect, the severity, onset, and duration. Economists also need to know who is potentially affected, specifying age, health status, income, and so on. And most important, they need to know how many people actually are at risk. The RfD/RfC and margin of exposure methods do not provide a quantitative estimate of risk below the point of departure, whether that would be a benchmark dose, or NOAEL/LOAEL, or LED10 for cancer. This is the focus of the colloquium.

In addition, the RfD/RfC addresses only critical effects from chronic exposure, not all the toxicological health effects as in the IRIS database, so they do not capture all the health benefits associated with a risk management action. The critical effects need to be related to adverse human health outcomes (liver disease versus liver weight change). Furthermore, in cancer assessment, more mechanistic information is used to emphasize reliance on precursor nonneoplastic response, rather than tumor data, and that should be reflected.

Emerging issues for future consideration include how more subtle effects can be valued as research provides new information on biomarker effects and susceptibility and mode of action at the cellular and molecular levels.

The overall goal for the colloquium is to explore possible approaches for quantifying risk below the POD, Vu concluded, specifically for effects presumed to have biological thresholds or for which mode of action information indicates that the dose-response curve at doses below observation could be nonlinear. Success means identifying some viable methods and approaches for the Agency to continue to develop for near-term use. She noted that EPA scientists would meet the following day to continue the discussion.

Background on Environmental Economics

Al McGartland, NCEE Director, described the mission of is center within the new Office of Policy, Economics and Innovation. His presentation slides appear at the end of this report as Appendix E. He welcomed all the participants, thanked those who put the colloquium together, and said he was optimistic about its chances for success. Even during the planning there was progress toward better understanding between economists and risk assessors. He called the colloquium a major milestone in bridging the gap between what economists need and what risk assessors are comfortable providing, and he invited people to participate in the work ahead.

Dr. McGartland briefly described the economic analysis guidelines due to be released, the paradigm behind benefit-cost analysis, and economists' methods and information needs. He also discussed as an example a case currently before the Supreme Court, in which the American Truckers Association has asked that EPA use benefit-cost analysis to set air quality standards.

Dr. McGartland described some history behind cost-benefit analysis for regulations. President Clinton's Executive Order 12866 is the current mandate. It requires a cost-benefit analysis of all major rules that have an effect on the economy of $100 million or more or that the Office of Management and Budget (OMB) believes are significant. In addition, the "Thompson language" is attached to the budget bill every year and requires OMB to report to Congress on the benefits and costs of all regulations. Each year OMB and EPA debate over how to capture the benefits that are not quantified, whether as zero or in some other way. In some cases, OMB assigns a value. The Safe Drinking Water Act and small business legislation in the Unfunded Mandates Reform Act require cost-benefit analysis, as do several bills currently pending on Capitol Hill.

Dr. McGartland emphasized that cost-benefit analysis merely provides a single input into the decisionmaking process but does not provide answers for what regulation or what option one should choose. Other considerations also need to be taken into account.

He explained Adam Smith's concept of "the invisible hand," which squeezes inefficient users of scarce resources out of the competitive market. He gave the example of a bad restaurant being driven out of business and a better restaurant taking its place. He said Smith was mostly correct, except with regard to market failures such as pollution. For example, clean air is a public good, not something that can be bought or sold. Cost-benefit analysis simulates how a private market would treat something like clean air. To do that, economists try to assess people's willingness to pay for a commodity such as clean air and determine whether the private market, if it could, would provide that good. Courts recognize the notion of that willingness to pay as a foundation for environmental damages, as in the *Exxon Valdez* oil spill.

Economists use two general approaches. The most frequently used is the damage function approach, which takes an interdisciplinary look at changes in emissions, exposure, environmental quality, and quantified risk reduction, then assigns a unit value per avoided effect. Another approach uses more indirect methods to value environmental improvements described in more general terms (for example, cleanup of the *Exxon Valdez* spill) without detailed enumeration of the specific improvements.

To get at unit values for reduced health effects, economists use several methodologies. The easiest, but least satisfactory, is cost-of-illness, which calculates the total cost of a disease, including hospital admissions, medical insurance, and so on. That cost is recognized as a lower bound but because it is a very hard number it can be very useful. Averting behavior teases out of people's observed behavior their willingness to pay. The hedonics approach takes market data and teases out the value of an attribute of an environment-related commodity. For example, regression techniques can be used to determine how much of the value of a home derives from the clean environment in which it is located. Wage-risk studies provide information on the wage premiums required for workers to assume risks on the job. To determine stated preferences through surveys, economists work with cognitive psychologists and scientists to establish the commodity to be valued, such as less exposure to toxics and therefore fewer health effects or

avoiding an environmental disaster like the *Exxon Valdez* spill. This approach is the most controversial but offers a way to get at values that are otherwise hard to capture.

In valuing a human life, economists look at a statistical human life but prefer to consider the value of small changes in the risk of mortality. Dr. McGartland mentioned the challenge in valuing the air toxics program as required by section 812 of the Clean Air Act because many variables are unknown. One approach is to try to value the safety EPA is providing the public through its air regulations. That makes the commodity of concern a broader notion of safety, rather than the risk of dying. Another approach is to think about the value of exceeding some threshold. He noted that without an ability to develop dose-response curves and quantify cases, it could become impossible to value anything.

Dr. McGartland discussed the two studies of Benefits and Costs of the Clean Air Act, required by section 812 of the act, which used a damage function approach to investigate many health effects that had relatively rich databases in the epidemiology literature. The first step was to quantify the cases after determining the change in emissions resulting from EPA regulations. That yielded a large number of cases to value. The numbers are controversial because they are so large, but the analysis showed that the Clean Air Act of 1970 prevented 23,000 premature mortalities annually, 20,000 chronic bronchitis cases, and similar numbers of hospitalizations for respiratory and cardiovascular disease. Then economists scoured the literature and determined a mean value for those health effects, in 1990 dollars, of $4.8 million, based on a fairly standardized approach for valuing risk reductions for mortality. Based on the 1990 amendments to the act, looking prospectively at the year 2010, the benefit in premature mortality alone is $100,000 million. Looking back to 1970, the ratio of benefits to cost was something on the order of 42 to 1, he said. It was somewhat less for the 1990 amendments but was still large.

The same methodology was used in preparing the economic analysis of the revised National Ambient Air Quality Standards for particulate matter and ozone. However, providing changes in margins of exposure would not mean the same thing to the public or economists who are charged with doing the cost-benefit work. Approaches are required to meet everyone's needs, grounded in the best science, but that can inform decisionmakers in the most realistic way possible about the trade-offs of different regulatory options.

Dr. McGartland concluded with the caveat that unquantified values are often counted as a zero. That complicates efforts to engage the public.

Discussion

One participant asked about the uncertainties associated with McGartland's numbers. He replied that they are very uncertain. The report to Congress included a whole chapter of Monte Carlo analysis. He noted there is a debate with OMB over what assumptions should go into a lower bound of the total benefits numbers. Often there can be overlap between choosing which epidemiological study to use and doing regressions, so great care must be taken in adding up

different kinds of categories to get a total benefit number. But, he added, the costs were always below even the lower bound of benefits.

Dose-Response-Based Distributional Analysis of Threshold Effects

Sandy Baird of The Baird Group and Lorenz Rhomberg of Gradient Corporation presented their findings. Their slides appear at the end of this report as Appendix F. Dr. Baird spoke first about methods developed collaboratively between toxicologists, risk analysts, and biostatisticians with a focus on the path from exposure to estimating the number of cases and providing a probabilistic estimate of those cases.

She said informed decisionmaking is complicated when it is based on RfDs. Uncertainty in the RfD is unknown; therefore, the protection from any particular RfD value is both unknown and inconsistent due to inconsistent data that go into the RfD. Decisions about levels of conservatism are actually risk management decisions, but they get intertwined with risk assessment during the process of creating an RfD and assigning uncertainty factors for each of the areas of extrapolation. At the end of the process it is unclear how much uncertainty there is. And there is no estimate of the risks of exposure at levels greater than the RfD.

Current models, the NOAEL and the benchmark dose, have the same underlying model. But many developmental studies show a 0 to 4.5 percent risk of a response rate at the level of the NOAEL, so the assumption that the NOAEL actually represents a dose where there is no effect is somewhat weak. Study quality also has a big impact on the ability to observe effects. Both of these methods apply a series of uncertainty factors. Dr. Baird and Dr. Rhomberg's talk focused only on the factors for animal-to-human extrapolation and accounting for sensitive human populations. Remaining factors representing subchronic to chronic, LOAEL to NOAEL, and data deficiency factors are related to not having as much data as one would like and can be accounted for probabilistically.

Dr. Baird showed a slide of three parallel dose-response curves depicting NOAELs from experimental animals, extrapolation using uncertainty factors to depict average humans, and a second extrapolation to depict sensitive humans. Dr. Baird and Dr. Rhomberg's model does not require the typical assumption that the NOAEL is below the population threshold after applying the uncertainty factors.

Dr. Baird built a dose-response model using the proposed methodology. It begins with animal experimental data in a dose-response curve. Uncertainty in the model is characterized as a full distribution that is carried through the model, unlike a benchmark dose where the lower confidence limit is the focus. The second step involves scaling, or centering, using whatever available methodology provides the most chemical-specific information to estimate the human equivalent concentration. The team chose the ED50 as the point of comparison between the animal and the human dose-response curves. In the third step there is a broadening of the distribution to account for the animal-to-human adjustment factor (AFA), adding in the

uncertainty in the accuracy of the scaling adjustment. Because variability is reduced as much as possible in animal experiments, the curve is flattened to depict greater heterogeneity in humans.

Since the curve now represents the distribution of human population thresholds, and depicts the uncertainty about the dose associated with levels of risk, a risk manager can decide what human population to protect (such as the 1/100 more sensitive individual) as well as how confident to be in the dose estimate. Finally, Dr. Baird provided the supporting equation (see Appendix F, p. 10).

Dr. Rhomberg described in more detail the theoretical and empirical evidence underlying the method, provided a brief case study of ethylene oxide that was recently completed using the framework, and summarized benefits of the strategy.

He walked backward through the methodology because the overheads were difficult to read. The goal of this methodology is a human dose-response curve showing proportions of the population expected to respond at different doses, based on the idea that each individual has a threshold and that as doses increase they exceed that threshold in an increasing percentage of the population. To express the uncertainty posed by extrapolating from an animal study, the dose-response curve also shows distributions around particular percentiles. The team chose to center the curve on the ED50 median effective dose as the most reliable point. That curve was achieved through centering, a theory of cross-species extrapolation that allows for the best available estimate of an equivalent toxicity between humans and animals, acknowledging the uncertainty in that process as well as from chemical to chemical. Backing up to the test animal, Dr. Rhomberg noted the animal dose-response curve is on a probit scale so it is really a matter of a probit curve being fit to the animal data. The experimental uncertainty of fitting the model is therefore represented, and carried through the whole process rather than just using a lower bound. He said the methodology tries to identify and account for all the elements of uncertainty that come from extrapolating, then tries to fill them with empirical distributions of how those extrapolation factors actually vary.

Dr. Rhomberg explained that instead of uncertainty factors (UFs), the methodology uses AFA and AFH, or human heterogeneity adjustment factor. AFA provides a central estimate of where a human response should be, given the animal ED50 response. That is separate from the characterization of uncertainty around that central estimate. AFH empirical data (developed and presented by Dale Hattis later in the colloquium) is used to show the variation among humans in sensitivity to toxicity.

Animal-to-human extrapolation can use various methodologies, such as allometric, RfC, and chemical specific, that are not dictated by the framework. But even the best extrapolation methodologies have uncertainties that must be taken into account.

Another question is how to characterize empirically the variation in scaling from one chemical to another. The team used various estimates based on distributions of relative potency. More

familiar in the cancer area, the concept considers the distributions of ratios such as potencies from epidemiological studies that are compared to animal studies. The availability of such ratios for noncancer endpoints is more limited. A study by Dr. Baird et al. (1996) of pesticide NOAEL ratios showed they tend to be lognormal distributions with a range of geometric standard deviations of 4.1 to 4.9. Rhomberg did a paper more recently with Scott Wolf on LD50 values, based on a larger database of more than 4,000 chemicals for some species-to-species comparisons, and found variations toward the lower end of the 2.5 to 6 range, mostly on the order of a GSD of 3. There are about 50 antineoplastic agents for which there are actual human data for maximum tolerated dose in a cancer therapeutic setting, compared with similar kinds of endpoints in animals over a series of several animals and humans, and those give GSDs in the same range (2.6 to 3.7).

Dr. Rhomberg showed one example of LD50 ratios for both guinea pigs and rabbits over about 3,000 chemicals. He noted the lognormal distribution is "more peaky" than most. He suggested the effect can be generalized, regardless of what species are used.

On human heterogeneity, the idea is that there is a tolerance distribution in the human population. Rather than using dose-response data from animals, which tend to be very heterogeneous due to study conditions, the methodology uses general human data on variation and sensitivity over lots of chemicals. In fact, while humans are slightly more variable, it is not as much as would be expected. That concept is described by a log-probit with a GSD that is derived from Dr. Hattis' empirical data.

Dr. Rhomberg also presented a case study applying the methodology to two developmental and two reproductive studies on ethylene oxide. The endpoints were postimplantation loss, a quantal endpoint, and fetal body weight, a continuous endpoint. The first step was to apply whatever advanced statistical dose-response modeling was available. That step captured some problems such as intralitter correlations and correlations among endpoints and covariants, like litter size, to characterize the full distribution of uncertainties and bring that through the population. Finally, sensitivity analysis demonstrated that controlling for litter size do not make much difference, while different GSD values associated with cross-species extrapolation do make a significant difference. Sensitivity analysis can suggest research strategies to focus on places where sharpening up information on a particular element would be the most beneficial.

He showed a graph of the results (see Appendix F, p. 17), depicting a fairly wide uncertainty distribution around the 0.1 percentile of the human sensitivity distributions, or the 1/1,000 risk level in humans. The mean was about 700 ppb.

To be very sure of the characterization of that percentile, one could choose an appropriate lower bound, such as 95 percent, and find the associated value, basically the ED01. This can be done for any percentile of distribution, so in cases where there are exposures at various levels in a population, uncertainties could be characterized for various dose levels and the associated

expected percentile. Therefore the method does allow for projections of a number of cases above or below the RfD with a full uncertainty characterization.

Issues that were not addressed included:

- Severity of effect
- Defining adverse effect
- Concordance of endpoints across species

But Dr. Rhomberg summarized the following benefits of the approach:

- Provides a distribution of the probability of a health impact occurring
- Estimates risks to specified sensitive subpopulations
- Quantitatively characterizes uncertainty in the risk estimates
- Determines the level of protection at the end of the process
- Estimates risk above and below the RfD
- Provides a framework for each component of extrapolation
- Allows for updating of components with chemical-specific data
- Allows for identification of components that contribute the greatest uncertainties so that resources can be allocated to reduce those uncertainties

Dr. Rhomberg concluded that the team believes the approach maximized the use of the available data in a framework that makes assumptions very transparent. It provides estimates of risk and uncertainty in those risks for sensitive human populations. He believes it is well suited for cost-benefit analysis that assesses the number of cases affected by a change in exposure.

Discussion

One participant asked whether the selection of the ED50 assumes that the log dose-response curves are parallel. Dr. Rhomberg replied that part of the purpose in choosing the ED50 is to avoid making that assumption. If they are not parallel, the point at which the comparison is made matters. The team selected the ED50 as the most appropriate point.

The participant asked whether the team had done a sensitivity analysis on the effects of choosing such a high effect level where variations would be expected to be smaller. She inquired whether the method introduces more uncertainty because of the extrapolation over several orders of magnitude. Dr. Rhomberg replied that the team has looked at the ED10 as an alternative. The magnitude of the effect depends on two things: how different the slopes are and how different from the ED50 the point is at which the comparison is made. Because the method is based on empirical distributions of relative toxicities, which are going to be near the middle of the distributions, he said the choice should not be too far from the ED50. He also noted that while humans have a broader distribution and shallower slope than animals, the difference is not as

great as might be expected because experimental animals were found to be more varied than expected.

Another participant asked whether two dose-response curves having the same ED50s and using allometric scaling under the methodology would result in the same answer.

Dr. Rhomberg answered yes, adding that there is a tradeoff regarding whether to start with animal dose-response curves or to take data directly on interindividual variability in humans. The former requires an assumption of interindividual variability; the latter is not specific to the endpoint and chemical of concern. The team chose to use general human heterogeneity. One alternative might be to expand the animal slope using a human flattening factor, but the team decided there was no real basis for doing that. However, it would be possible to apply a correction factor to the general steepness of animal dose-response curves to obtain a general steepness for human dose-response curves. He noted that while the method makes a general accounting of all the sources of uncertainty, it does not yet address the fact that, from chemical to chemical, the degree of human heterogeneity probably varies. That can be captured in a distribution based on Dr. Hattis' data, but it has not yet been done.

The participant inquired whether using the ED50 instead of the ED10 loses some of the heterogeneity of the animal data.

Dr. Baird replied that the heterogeneity in the animal population is carried through in the distribution of the stochastic experimental uncertainty. A better-quality experiment provides a narrower distribution on that stochastic uncertainty. So, using two chemicals, one with a better quality study than the other, would show a narrow distribution of the uncertainty in that estimate that would be carried through to the end.

Dr. Hattis remarked that his presentation would show there tends to be more variability for less severe effects, and for some kinds of organ systems. So to the degree that the effects of concern can be subcategorized, there can be some clues to adjust the human dose-response relationship accordingly. Dr. Rhomberg added that the team tried to design an approach that provides a natural place for adding refinements as they are developed.

Characterizing Risks Above the Reference Dose

Paul Price of Ogden Environmental and Energy Services reported on work done by his firm, in which the reference dose was put into the dose-response framework and the traditional uncertainty factors used in deriving the reference dose inside that framework were defined. He also proposed two approaches for defining risks above the RfD and discussed the implications for assessing carcinogenic risks. His presentation slides appear at the end of this report as Appendix G.

The current approach for noncarcinogenic risk assessment calls for setting a permitted dose, but that is never associated with a likelihood of response for doses either above or below it. As a result, use of the risk tools that are associated with this permitted doseùthe risk cup, the MOE, and the hazard quotientùprovide no guidance for the benefits associated with reduction. His presentation was the result of work done by his firm in collaboration with EPA and TERA under a cooperative research and development agreement to investigate the uncertainty in the variation in the RfD and in assessments of noncancer risk. The work resulted in four publications and has been used to develop quantitative estimates of noncancer risks for both PCBs and mercury.

Mr. Price noted that in his approach the reference dose is a technical finding rather than the product of a political or social process. Before talking about risks above the RfD, it is important to establish a framework that clearly states variations versus uncertainties. He suggested that framework would be absolutely consistent with Drs. Baird and Rhomberg's. Variation must be stated in terms of differences in the relative sensitivity of individuals, usually a fraction of the population that will respond at a certain dose (i.e., the dose-response curve) or the amount of interindividual variation that contributes to sensitivity differences. In contrast, uncertainty comes from either measurement error, or to represent the extrapolation from the animal model to human. He suggested that toxicological criteria such as the RfDs are best understood by stating they only have uncertainty.

Mr. Price's approach postulates that each individual has a threshold and that the distribution of the thresholds goes to zero, allowing for discussion of a population threshold. He also addressed the implications of a situation where there is no population threshold. He noted that the differences between individual thresholds and population thresholds are clearly a matter of variation.

Uncertainty comes in because the distribution of individuals' thresholds cannot be measured. Therefore it must be either modeled mathematically or based upon animal surrogates or other indirect methods for determining the shape of the curve in humans and whether it has a population threshold. That results in a true but unknown dose-response curve with confidence limits around the estimate. The lower confidence limit of the dose that causes a zero response is the lower confidence of the population threshold. That falls on the curve of the upper confidence limit of the response rate.

To provide a basis for the mathematical approaches for calculating risks above the RfD, Mr. Price defined the RfD as an estimate of the lower confidence limit on the estimate of population thresholds in humans rather than as an estimate of the population threshold. In some cases where the RfD does not correspond to zero risk, it is actually the lower confidence limit of some finite but very low level of risk.

He suggested that taking the standard equation for the reference dose of the NOAEL, dividing it by the various uncertainty factors, and replacing them with distributions does not result in the distribution of the RfD. Rather, the result is the confidence function for the estimate of either the

population threshold or the estimate of the dose that causes a very low level of risk. At some point on the lower confidence level of this distribution is the actual RfD. By substituting distributions for each of the terms in the RfD equation, the result is not uncertainty in the RfD but uncertainty in the population threshold or some EDR, where R is a very small number, and the RfD is best understood as a lower value taken off this distribution.

Mr. Price divided the traditional uncertainty factors into three categories. The primary factors, of inter- and intraspecies uncertainty factors, must be used to go from a NOAEL in an ideal data set for animals to the RfD. The secondary factors are all the other factors that are necessary to try to estimate that NOAEL in an appropriate data set and reflect data limitations. The third category is all the modifying factors, such as FQPA, which are adjustments in the reference dose that reflect other concerns.

He suggested that the intrafactor is best understood as another type of interspecies uncertainty factor. The classic interspecies uncertainty factor refers to the average differences between the sensitivity of two species. Starting at the ED50 provides a useful way of separating the differences between an average member of each of two species that can be well understood by pharmacokinetics or readily measured and the differences in the slopes or the dispersion around that mean between two species. He proposed understanding the interspecies factors in terms of differences between the average sensitivity of two species, typically the animal and the human, and the intraspecies as reflecting differences between the dispersion or slope of the dose-response curve in the animal model as compared with the human model.

Mr. Price showed a graph depicting that concept, proposing that the UFA be understood as either a scaling or a simple displacement of the dose-response curve in the animal to arrive at a dose-response curve in humans showing a log dose response. The sensitive individual factor is used to change the shape of the dose-response curve, leading to a flatter dose response in humans. It also leads to an estimate of the RfD that is a function of moving from some sort of NOAEL or benchmark that is a measure of the low response rate of the animal divided by the UFA or UFH to come up with an estimate of the RfD. Noting that the approach is rather simplistic, he suggested it provides a firm basis for estimating risks above the RfD and a firm starting point to talk about traditional complexities such as the true shape of the dose-response curve in humans, uncertainty in estimates, and variation.

He then proposed two approaches. The first used basic algebra, based on the prior assumptions. The second, assuming the shape of the dose-response curve in animals has no particular relevance to humans, took from the animal study only some measure of the dispersion, or very general description of the dose-response curve.

Under the first approach, inter- and intraindividual uncertainty factors move the center through a simple displacement and by changing the slope. Simple algebra produces an equation,

$$EDRh = (ED50a/UFA)(EDRa/ED50a)(1-(\log UFH/\log(ED0a/ED50a)),$$

which relates the estimate of the dose causing a response R in humans as a function of the ED50 in animals (ED50a), the size of the interspecies uncertainty factor, the dose causing a response R in animals and the intraspecies uncertainty factor or UFH, as well as the estimate of a dose causing zero or extremely small response. The equation comes from the requirement for one scaling factor to move the entire curve over and another associated with intraspecies to move the lower portion of the curve over by an amount equal to the uncertainty factor.

This equation would be used by putting in a mathematical distribution of an animal curve to map over the shape of the curve. He showed an example using factors of 10 for inter- and intraspecies. At the 50 percent response level, the curve moves over by a factor of 10, while at the estimated threshold, it moves over by a factor of 100. In between, the move is a linear function of the difference in response between zero and 50 percent. While very simple, the approach builds on the existing definitions of the uncertainty factors. The equation allows inclusion of a distribution for the uncertainty factors that reflect estimates of the threshold or the ED50 that can be carried through to determine uncertainty confidence limits, the uncertainty around the predictions of the dose response in humans. A major drawback, however, is that it requires the assumption that the shape of the dose-response curve in the animal is relevant for humans. The approach does require an estimate of the threshold in the test species, which can be derived either from a mathematical consideration, such as some policy that starts with the benchmark dose and divides by some factor, or based on pharmacokinetic or other mechanistic arguments. To do a meaningful Monte Carlo version, he said, the selected values must be correlated.

A more minimalist model takes a couple of points on the animal dose-response curve and simply draws a straight line between them, using the same mathematical assumptions to extrapolate those data points over to humans. That model has the advantage of not requiring the assumption that a dose-response curve in animals predicts that in humans. Taking data in the animal species, Mr. Price showed the actual curve that may truly exist or may be predicted from models. Then, the value of the ED50 and the RfD are moved over, generating a "hockey stick equation" that is used to predict risk where the doses below the population threshold give zero response and the doses above it just follow a straight line from the estimate of the population response up to the estimate of the ED50.

The equations for doing this are very simple, based on available data (NOAEL, the ED50, the uncertainty factors). Mr. Price's team published a paper applying the method to four compounds for which data were available for estimating the NOAEL and the ED50. The paper used a distribution for the uncertainty factors to estimate the median values of the dose response and the upper confidence limit. Mr. Price showed a slide illustrating the results for the median values, with hazard indices from zero up to 10 times the hazard index of 10 (corresponding to a dose 10 times the RfD). The model gives a best estimate prediction that no risk occurs until the dose increases 10-fold, then increases proportionately after that. At 50 times the RfD, estimates of risk go anywhere from about 8 percent to about 30 percent of the population responding. He

showed another slide of the same data showing the upper confidence limits. The risks at 10 times the RfD were therefore not zero as the earlier slide showed, but about 4 percent to 15 percent, reflecting differences in the magnitude of the uncertainty factors.

For cancer risk determination, Mr. Price noted that either of the two approaches can be used to estimate response. Obvious modifications include replacing the interspecies factors traditionally used in noncancer assessments with assumptions that come from the cancer tradition of extrapolation based on body weight to the two-thirds power, rather than a factor of 10. He suggested that Dr. Hattis' data on the traditional variability may be a better basis for developing estimates of the magnitude of the intra-individual factor. And the approach is amenable to calculating uncertainty using the Monte Carlo model.

In contrast to the Baird and Rhomberg approach, Mr. Price's approach requires choosing a dose believed to be associated with the threshold and the confidence limits around that dose.

In summary, the approach provides a basis for beginning to talk about deriving quantitative estimates. While intended to reflect the existing RfD framework and traditional uncertainty factors, it can be used to look at other issues. It is not limited to noncancer, but can be applied to any endpoint where a population threshold is believed to exist above zero. And it allows separate modeling of variation uncertainty. He concluded that using all the equations and forcing the threshold to go to zero produces something that looks very much like the existing cancer quantitative risk numbers.

Discussion

One participant asked whether using the ED50 would work as well for asymmetrical distributions as for ones where the median and mean are the same. Mr. Price said the analogy would still hold, but the question gets into difficulties in describing multimodal measures of dispersion, rather than the simple lognormal distribution described by GSD. The complexity of that result requires many more data to suggest how humans actually vary and how sensitive subpopulations differ from typical individuals. But although the math would be more complex, he said, it would not be difficult to do inside the framework.

Another participant asked whether, in the approaches presented by both teams, the population distributions account for an effect that might be specific to some segment of the population, such as a developmental effect. Mr. Price said the reference dose approach tries to establish a dose that protects the entire population, and thus considers endpoints that may affect only a portion of the population. Addressing this issue comes not in the toxicology portion but in the exposure portion of the benefits analysis, where the goal should be to describe the distribution of doses received by the portion of the population to which the RfD applies. It also suggests that benefits might be missed by only looking at reproductive effects of dose among women in child-bearing years. So it may be desirable to have separate dose-response curves for the general population to

investigate the benefits associated with reducing their exposures as well. That would have to be done by starting over from scratch with the toxicological database, not just with the RfD.

Dr. Rhomberg concurred, stating that in his team's case the effect was aimed at a particular portion of the population, assuming that is taken into account in the exposure analysis and in identifying the number of people who are subject to those risks. It is more difficult to start identifying that population based on some sensitivity, and not just qualitatively, because then the calculations of variation and sensitivity are confounded with the definition of the population at risk.

He noted that the example raises an issue usually considered under sensitive endpoints. Sensitive endpoints can be identified under the traditional RfC/RfD framework, and as long as everybody is protected from that, they are protected from everything else. But it is not so clear in thinking about a risk above the RfD. Moreover, many of the endpoints are graded by severity. Dr. Rhomberg said that without further modification the team's method really only predicts a low grade of severity and does not allow for a higher severity among people who are particularly sensitive.

The discussant made the point that doing benefits analysis may require looking at something broader than the typical RfD.

Mr. Price noted that as the critical factor on the RfD increases to an ever-larger fraction of the population, there is potential for other noncritical effects that occur at higher doses but are of far greater concern, such as frank effect levels. So in looking at analysis showing only a 20 percent chance of suppression of liver weight, there may be a 3 percent chance of a truly frank effect, the avoidance of which would be associated with a great economic benefit. Approaches that start with the RfD and the critical factor would have to be modified to address that.

One discussant noted that an advantage of the Baird-Rhomberg approach was in recognizing that there can be some residual risk at the RfD. She suggested that perhaps Mr. Price's approach could be extended to model some residual risk at and below the RfD. Mr. Price agreed, saying that comment captured one of the main critiques of the published paper. He said allowing for a finite level of risk at the reference dose would not be hard to introduce into the equation; in fact the paper offers a method for doing that. But he said his approach does not accomplish that as elegantly as that proposed by Dr. Baird and Dr. Rhomberg.

Expected Values of Population Dose-Response Relationships

Dale Hattis of Clark University spoke about human risk. His presentation slides appear at the end of this report as Appendix H. Many different parameters have been measured in humans that are helpful for quantifying risk as a function of dose. The database of pharmacokinetic and pharmacodynamic effects includes mostly pharmaceutical data, with very few environmental

data. The parameters measured cover different portions of the pathway from external exposure to internal response.

There are a number of different routes to causation and quantification of health effects. Cases where direct human epidemiological observations can be made do not require as much risk assessment, just direct measurement tools and controls for confounders. Indirect projections also can be made of possible effects as a function of changes in various intermediate parameters, assuming that whatever is causing those changes is having a parallel impact on quantal effects like mortality. This is a relatively underdeveloped area.

Dr. Hattis looked at population distributions of individual susceptibility to effects that occur as a result of overwhelming homeostatic systems. He divided the pathway from external exposure to internal response, identifying variabilities in

- contact rate that can affect the projection, such as a distribution of water or food consumption,
- variability in uptake or absorption, per unit intake or contact rate,
- general systemic availability, net of first pass elimination,
- dilution via distribution volume, and
- systemic elimination or clearance.

Those variabilities are part of pharmacokinetics. Pharmacodynamics includes

- how much an internal physiological parameter (e.g., FEV1) is changed per unit of internal concentration, and
- how much change in that physiological parameter is required to achieve some effect of concern.

The database currently contains 443 data sets, each measuring variability in a particular relevant pharmacokinetic or pharmacodynamic parameter for a particular chemical. The bulk of the database provides pharmacokinetic data, such as half-lives and volumes of distribution. There is now also a respectable body of pharmacodynamic information, about 89 data sets. While the database has been expanding to include some data from children under 12, there are not as many data of as high quality for children.

For the distributions of human inter-individual variability of different parameters, Dr. Hattis assumed lognormality and showed some plots to support that assumption. He combined the variability from multiple causal steps by adding together the lognormal variances associated with each step. Different chemicals have different amounts of variability. Plotting the measures of human lognormal variability for particular parameter types indicates that the log(GSD)'s themselves are approximately lognormally distributed. But there is still a question about how much of the differences seen in individual variability observations is due chemical-to-chemical variation and how much is due to measurement error, which tends to inflate the observed

variability. That must be deflated to get as realistic a measure as possible to obtain a central estimate of the amount of variation and therefore the amount of real uncertainty in the slope of the population dose-response relationship in people.

The basic methodology for assessing expected values, or arithmetic mean risks averaged over the relevant uncertainties and capturing uncertainty as fairly as possible, requires an explicit treatment of uncertainty. This is because these distributions tend to be highly skewed. And the expected value or mean of the distribution is generally larger than the median, so fairly capturing uncertainties is important.

Dr. Hattis' process uses the human database to make a central estimate of overall lognormal variability from the observed variances associated with various causal steps. Depending on the kind of exposure data it starts with, the route of exposure, the effect, and the degree of severity, the variabilities associated with those causal steps are added together. Then, based on the available observations, one can determine the lognormal uncertainty in the log(GSD)'s themselves and reduce the inflating influence of statistical sampling error on the observed spread of those values. Next, Dr. Hattis sampled repeatedly from the assessed lognormal distribution of log GSDs. (GSD is the standard deviation of the logs to base 10 of the individual parameter values, or the log to the base 10 of the geometric standard deviation.) He calculated the arithmetic average of risk for people exposed at various fractions of a human ED05 level, although with some extra standard deviations it can easily be done from an ED50. It can be done in a Monte Carlo format or as a simple spreadsheet. Dr. Hattis said that the surprising bottom line is that he can summarize the expected value as simple power law functions, basically as graphs on log-log plots that look relatively straight but have different slopes depending on the variability and the spread of the log GSD variability observations that are chosen to represent the uncertainties.

The approach does not address the underrepresentation of children, the elderly, and sick people in the studies, which is likely to result in some understatement of variability. On the other hand, some measurement errors in the primary variability observations act in the opposite direction. That can be analyzed, but Dr. Hattis said he has not yet done so. Another difficulty lies in the fact that the drugs studied might not perfectly represent the environmental or occupational chemicals of interest to EPA.

Dr. Hattis explained the meaning of the log GSD numbers relative to the 10-fold baseline assumption. Using a 3.3 standard deviation range from a 5th percentile to a 95th percentile, the number of standard deviations is similar to what is needed to go from an ED05 to an ED10-5. In that range, the 10-fold baseline corresponds to a log GSD of about 0.3. Below 0.3, the 10-fold factor allows for a projection from a low or no effect level to an incidence of effect that is potentially socially tolerable on an RfD basis. Going much beyond that would require a larger reduction of dose than the 10-fold factor provides in order to achieve a comparable reduction in expected risk.

Dr. Hattis showed a slide depicting a set of raw, unweighted pharmacokinetic results in the general range of 0.1 to 0.2, with upper 90 percent confidence levels going up to the order of 0.3. Most of the time, pharmacokinetics alone fall within the 0.3 level. But the pharmacodynamic portion of the database is usually larger. That can result in variabilities of very substantial amounts, such as in cases like local contact site parameter change, where the central value is around 0.6. Systemic effects from external administration are often less than that, with central values on the order of 0.2, but with 90th percentile values in the range of 0.6-0.8, in rare cases approaching 1.

He showed a few actual distributions as examples. One was a probit plot with a Log(GSD) on the order of one, of the fraction of people who suffer a particular kind of skin hypersensitivity as a function of log of the chromium concentration applied on the skin. Another was a plot of the distribution of concentrations of the chemical methacholine that cause a 20 percent decrease in FEV1 in smokers with mild to moderate air flow obstruction. The data set included 5,000 people and provided a fairly good lognormal distribution for the four different concentration points on the curve. Another big data set for histamine showed the same sort of distributions, but a log GSD on the order of 0.6. He also showed a plot of aggregate log GSD distributions, basically log GSDs for pharmacological half-lives in adults and children, that was basically lognormal.

With pharmacokinetics, Dr. Hattis concluded, there is not too much overall variability. Results may not be perfectly lognormal, but the approach is feasible for trying to fairly capture some of the uncertainty.

Discussion

One participant inquired how widely applicable the approach is, whether it also works with more subtle effects. Dr. Hattis replied that it works when individual thresholds for a response can be defined, noting that some individuals will always be more sensitive. It can be used on more subtle effects and with background. But a population threshold should not be expected if there is background dose.

Another discussant asked whether the less severe, reversible effects that were measured, such as moderate nose irritation, eye irritation, and skin rash, were subjectively reported. Dr. Hattis noted that this is a worry and reporting error variability needs to be teased out.

One participant pointed out that phase two drug trials routinely include genetic screening, which would narrow the variability found. He asked whether data are available to determine the impact. Dr. Hattis noted that the Food and Drug Administration drug database is extensive and completely secret. He said a cooperative effort between EPA and FDA, providing adequate privacy protection, could be of great value.

Risk-Based Reference Doses

Dave Gaylor of Sciences International presented work on risk-based reference doses done in collaboration with Dr. Ralph Kodell at the National Center for Toxicological Research and FDA. His presentation slides appear at the end of this report as Appendix I.

Current noncancer safety assessments can be based on a NOAEL or LOAEL, but do not characterize the risk. He suggested starting instead with the benchmark dose, which provides an estimate of the risk.

The median of the overall distribution of the product uncertainty factors can be found using the individual median of the individual subcomponents. The overall standard deviation then is just simply the square root of the sum of the variances associated with each of the uncertainty factors.

Dr. Gaylor showed a formula illustrating how to translate estimates of the median values and standard deviations for effects into percentiles. He referred to a 1983 paper by Dourson and Stara, reviewing published literature on a few dozen chemicals having results for more than one strain of a species. They used that paper to obtain a measure of interindividual variability to estimate variation in a human population. The ratio of the individual values for a strain compared with the overall median provides a measure of interindividual uncertainty.

Using an uncertainty factor of 10 for human variability covered about 92 percent of the chemicals in the database. To cover 95 percent required a factor of 15. Dr. Gaylor does not advocate changing the default value because additional conservatism results when several uncertainty factors are multiplied together. But with a single uncertainty factor, 10 generally is not large enough to give a high confidence of covering a large percentage of the chemicals. Calculated from Dourson and Stara's set, the standard deviation for the random variable, the uncertainty factor for human variability log base e, is 1.64. In terms of base 10, that is a standard deviation of about 0.7.

Starting with a benchmark dose of 10 percent and assuming a lognormal distribution for human variability, it is possible to plot various specified levels of risk, such as 1/10, 1/100, or 1/1,000. Stated another way, given a reference dose, one can determine the number of standard deviations from the benchmark dose and estimate what the risk is. This gives a procedure for estimating risk at the reference dose or for setting a reference dose associated with a specified level of risk. Two things are required to do this: a benchmark dose as a point of departure (Dr. Gaylor started from an ED10), and an estimate of the standard deviation. Dr. Gaylor used a standard deviation base e for variability among humans of 1.7 (a middle ground value) in cases where nothing is known about what kind of variation to expect for a particular endpoint or a class of chemical.

On Dr. Gaylor's chart, the risk of 1/10 was right at the benchmark dose of 10 percent, corresponding to 1.28 standard deviations below the median. A risk of 1/10,000 was 3.72 standard deviations below the median or about a factor of 63 below the benchmark dose for 10.

Ordinarily starting with a benchmark dose with a 10 percent adverse effect level, and applying a factor of 10 for going from an effect level to a low level and another factor of 10 for human variability, results in a factor of 100. At a risk level of 1/10,000, a factor of 60 would be sufficient if this estimate of standard deviation is appropriate. A risk of 1/1,000,000 requires a factor of 365 below the benchmark dose.

To provide for other uncertainty, another factor of 10 can be added for extrapolating from animals to humans, and another to extrapolate from subchronic data to chronic effects. It is possible to increase the confidence of this risk estimate by including additional uncertainty factors or starting from a lower confidence limit on the ED10.

Taking a benchmark dose of 10 percent and dividing it by 100, with this estimate of variability, gives an estimated risk of about 3/100,000. It is based only on some estimate of the standard deviation of variability expected among individuals. Given information about a particular endpoint, such as mortality, and estimates of standard deviation, this procedure could be improved by using that standard deviation. With information about a certain class of chemicals, a more specific estimate of the standard deviation could be used.

Dr. Gaylor noted that all the presentations rely on a log probit and assume a lognormal distribution to obtain an estimate of risk that can be observed in an animal study, whether at the 10 or 50 percent level. He noted that the procedure is basically what Mantel and Bryan published in 1961 for cancer data. Looking at cancer data, they suggested starting with a point of departure (they did not call it a benchmark dose) of a 1 percent risk level. They used a slope of 1 probit over a factor of 10 that is shallow and conservative. Dr. Gaylor stated that the standard deviation slope of 1.35 over a factor of 10 he used is not much different. The same approach can be taken now, i.e., pick a shallow slope to be conservative based on the best estimate for a particular endpoint or class of chemicals.

To estimate risk at reference doses requires starting with a specified benchmark dose rather than LOAELs or NOAELs. But it does not require extrapolating a dose response below a benchmark dose. It does require an estimate of the standard deviation for interindividual variation. With these approaches, risk can be estimated at a reference dose, above it, or below it. Or, dose can be calculated based on a specified risk for a certain endpoint. Confidence limits are derived from those on the benchmark dose. Other uncertainty factors can be introduced. Hence, it is possible to estimate risk at reference doses, or estimate a dose associated with a risk given an estimate of human variability.

Use of Categorical Regression to Characterize Risk Above the RfD

Lynne Haber of TERA presented categorical regression as a method for calculating risk above the RfD. Her presentation slides appear at the end of this report as Appendix J. The method was developed at EPA's NCEA-Cincinnati office.

In categorical regression, the toxicologist makes a judgment for each dose as to what the severity level is. Several different severity ratings have been used, such as no effects, minimal, moderate, and extreme. Some of the morning presentations did not address different severities. Categorical regression can include multiple studies and model incidence data, continuous data (such as the mean), as well as qualitative data, such as observations of liver necrosis at a given dose level.

A dose-response curve is fit to the data using the different severity levels. Earlier studies used a logistic regression. Current modeling abilities allow other mathematical forms to be fit also. The toxicologist can evaluate (or stratify) the data separately by the endpoint and by the species to see whether data from different endpoints or species can appropriately be combined or whether there are differences that need to be taken into account.

The results of the modeling are judged graphically as well as by the data quality. Several statistical tests can be used.

There are several advantages to categorical regression. The data requirements are much less rigorous than those for benchmark dose modeling, so categorical regression can be applied when the data are not sufficient to calculate an ED10. The modeling approach also addresses an increasing severity of effect with increasing dose. It is possible to combine different studies using this analytical technique, and it is also possible to take duration into account as one of the parameters used in the modeling.

Limitations include the need to account for animal-to-human extrapolation and a loss of variability information in categorizing a group based on the mean response. But there are methods to take that into account.

Dr. Haber provided an example showing data from clinical studies of aldicarb exposure with human volunteers. As dose increased, the severity effect increased, as did the percentage of people affected. The frank effect level was defined based on nausea, vomiting, and lightheadedness, different from what is typically thought of as a frank effect. The data were modeled on a graph of exposure versus probability of response at that dose. As the dose increased, the response was higher. After a certain level, the probability of an adverse effect decreased because the probability of a frank effect increased.

Another slide showed the calculated probabilities of the adverse or frank effect or higher. As expected, there was minimal or no effect predicted at the best estimate of the RfD. The upper confidence limit was 10-5. Based strictly on curve fitting, Dr. Haber pointed out predicted probabilities and upper confidence limits of having the probability of an effect based on the data.

She noted some advantages and strengths of this example. Human data were used so there was no extrapolation from animals. The data did show several no effect levels and no adverse effect levels, so there was not much extrapolation below the data. A weakness was small group sizes, which raises questions about how to account for sensitive populations. There are two

possibilities: (1) that the sensitive populations would be part of the same dose-response curve so that the dose response for those sampled adequately characterized the entire population, or (2) to have a separate curve of the sensitive population that differs qualitatively from the people who are studied. That case requires some sort of uncertainty factor to predict the risk in the sensitive population. But to predict the overall general population risk, one needs also to know what percentage of the total population is composed of that sensitive group.

Dr. Haber showed a second example comparing pesticides. That work was done to assist a risk manager in prioritizing when the risk cup is exceeded for several different pesticides. She said prioritization and the relative risk between the different chemicals is as much an issue as predicting the absolute risk. The work was done with data from animal studies and included both incidence and continuous data (where dose groups were categorized by severity of effect). The plot is of the log of the dose over the RfD, so if the dose is normalized to the RfD for the chemical, the plot represents the probability that the dose level is an adverse effect level. The graph showed three cholinesterase inhibitors, modeling the same effect for three chemicals that act via the same mode of action. For disulfoton, the lower risk is predicted at a given multiple of the RfD. This chemical had a larger uncertainty factor because the RfD was calculated from a LOAEL instead of a NOAEL, but calculating a benchmark dose could help refine the analysis. But it does give some sense of the relative risk in that the first two chemicals have very similar risks at a given dose relative to the RfD.

The plot shows a probability of having an adverse effect, but could also be viewed as equivalent to a population risk, assuming 100 percent are affected if the dose is an adverse effect level.

There are a number of issues in translating the relative risk of different chemicals to the expected human population risk at the dose relative to the RfD. The slope of the animal dose-response curve is expected to be steeper than the slope of the human dose-response curve. The doses that were modeled used a body weight to the two-thirds adjustment, so there has been a toxicokinetic adjustment in the dose-response curve. However, there still are toxicodynamic differences between humans and animals that were not taken into account, and the slope of the dose-response curve reflects the variability in response of the subject population. Because animals have much less variability than humans the slope for human data would be shallower.

Dr. Haber showed a slide depicting a model for two chemicals with two different modes of action: EPTC, with a very steep dose-response curve, and lindane, with a much shallower curve. Unlike with the cholinesterase inhibitors, there is quite a difference between predictions for the two chemicals. In the low dose range, the prediction for lindane is higher than that for EPTC, whereas in the higher dose range the predicted risk is lower.

When extrapolating from animals to humans, there are several issues to take into account in using categorical regression to predict the low-dose risk. One is sensitive populations. Does the dose-response curve include the sensitive populations if human data are available? If extrapolating from animal data, how are differences in the shape of the dose-response curve

between animals and humans accounted for? How do we account for the use of uncertainty factors? These questions were not addressed by the study but are becoming more of a concern. Another is model dependence, which increases with the distance from the range of the data. This has not been considered in great detail, but that is also a concern in extrapolating to lower doses. Do we force the model to go to zero at the RfD because toxicologists may expect that the RfD is a subthreshold dose, or do we just let it go mathematically wherever it goes because the RfD can be considered as an expected probability of response? Other issues include what data are chosen for modeling, what are the criteria for excluding studies due to low quality, what are the rules for assigning severity categories, what are the rules for combining studies, what constitutes an acceptable model, and how are results interpreted.

Other advantages are inclusion of all the useful data in the quantitative analysis, the possibility of meta-analysis, ability to take duration into account, and providing a consistent basis for calculating the risk above the RfD. Overall, categorical regression modeling is not as well developed for calculating the risk above the RfD as some of the other methods presented, but it has utility.

Dr. Haber addressed several more general issues raised in the earlier talks. She noted that intra- and interspecies factors can be used to characterize variability or variation, which can be a known quantity, although there may be uncertainty involved. LOAEL to NOAEL extrapolation can be dealt with given the data for a benchmark dose. But the subchronic to chronic and database uncertainty factors address uncertainty. And the proper numbers for those uncertainty factors cannot be known, only estimated (otherwise they are no longer uncertainty factors). In contrast, uncertainty factors that address variation can be broken down into, for instance, toxicokinetic and toxicodynamic aspects. There is a movement under way to use compound-specific data to replace the default uncertainty factors. The data may be specific to the chemical, the class of chemical, or the mode of action. It is important to consider, as the chemical specific-data become available to enhance various methods of analysis, how such data would be used in evaluating the predicted variability in response for a given chemical.

She also noted that not all uncertainty factors are equal. Drs. Baird and Rhomberg talked about using an entire database to characterize the uncertainty factor, while Mr. Price uses first principles about the sort of data that go into an uncertainty factor. At present, none of these approaches take into account explicitly the fact that the year of the assessment affects what is meant by a given uncertainty factor. During the early 1980s, an uncertainty factor of 10 was the default. Now, it means experienced risk assessors have made a judgment that the data are insufficient to reduce or modify the default. In the future, 10 may be a true chemical-specific adjustment factor. That meaning will affect the appropriate distribution to be used for that uncertainty factor. How that is taken into account will be important as compound-specific adjustment factors become available.

Discussion

One participant asked what effect was considered in the aldicarb study. Dr. Haber replied that all the test subject data were evaluated and categorized by severity. Most of the categorization was based on plasma cholinesterase inhibition, and alternative analyses were done based on whether certain levels of inhibition were considered adverse. The frank effects included clinical effects such as nausea and lightheadedness.

The participant asked whether an uncertainty factor was used for childhood sensitivity. Dr. Haber said the initial RfD was done with a basic uncertainty factor of 10. For the extrapolation down to low doses, at the time the work was done (in the mid-1990s), it was assumed that the dose response for humans adequately characterized the dose response for the whole population, so no other adjustment was made. That assumption could be looked at more carefully. There are many ways in which the work could be enhanced.

Risks Between the LOAEL and the RfD/RfC: A Minimalist's Approach

Resha Putzrath of Georgetown Risk Group began with a suggestion that models should be made simple, but noted that Mr. Price's analysis was more minimalist than her own. Her presentation slides appear at the end of this report as Appendix K.

She proposed three possible questions to be answered. Do we like the current RfD/RfC method at all? She suggested a strong case could be made for starting afresh with all the data, information, and models produced over the last 25 years rather than trying to improve various uncertainty factors or estimates of NOAELs, LOAELs, or thresholds. Assuming use of RfD/RfC methods, how can they be expanded to the risk of exposures above the RfD? How can the current point estimates be made more amenable for combination and comparison with data that tend to have ranges and distributions? She suggested it is hard to answer all three questions at the same time.

Assuming some confidence in the RfD/RfC method and looking at what can be done above it, Dr. Putzrath urged thinking about risk estimates between the RfD/RfC and NOAEL and between NOAEL and LOAEL. She also suggested looking at carcinogens with curvilinear low-dose dose-response curves.

Dr. Putzrath mentioned problems she has had with margin of exposure. She noted there are at least two distinct definitions of margin of exposure within EPA. One is used for noncancer risk assessments that assume a threshold but absolutely nothing about the shape of the dose-response curve. The second is, for carcinogens, assessments that assume there is no threshold and the dose-response curves are highly curvilinear. Although the definitions in terms of a point of departure over an expected exposure are the same, the mathematical implications are quite different and should not be described using the same term. In addition, she said, the inherent

initial assumption with margin of exposure is that it is somehow linearly proportional to risk, but that can be proven false.

Taking carcinogens, Dr. Putzrath said her minimalistic solution assumes the dose-response curves are not curvilinear. Claiming that they are requires a lot of data, which provides a dose-response curve that can be used to calculate an upper and lower bound. Without a lot of data, or assuming the lower response regions differ from the known dose-response curve, she suggested the method is not critical if one is not using the dose-response curve. But new methods often focus only on the upper bound of the risk curve, and that creates a problem in trying to make the true dose-response curve and its upper bound continuous functions that intersect at the origin. More importantly, she said, the evaluations of upper bound risks deal with individual chemicals. That focus ignores a maximum likelihood curve (best estimate of what is actually going on), which is required in order to combine evaluations and recalculate upper bounds for mixtures. That is a problem with the 1996 Proposed Guidelines for Carcinogen Risk Assessment.

It is important to note that, while analyses can be done for generic distributions, chemical-specific distributions, groups of chemicals, or same methods of action, they may not be terribly useful. She proposed instead looking at whether the data can be used to answer the questions simply.

Assuming existing RfD/RfC methods have value, the questions become: Is the dose-response curve, in the low dose and low response region, shallow or steep? How close is the NOAEL to the RfD? Is the exposure of interest closer to the RfD or the NOAEL or the LOAEL? That is one measure of the accuracy that is needed. The bottom line is, what are the consequences of a wrong answer or suboptimal decision? How likely are we to over- or underestimate the risk? Slightly underestimating a fairly minor effect is not the same thing as slightly underestimating a lethal effect. She also noted there are opportunity costs of misallocating resources.

Dr. Putzrath suggested a simplest case. Taking a relatively steep dose-response curve and an RfD that is a distance from the NOAEL (or LED10 or LEDx), the decision about how much accuracy is required should differ depending upon whether the exposure is e, e-prime, or e-doubleprime. In particular, if RfDs are meaningful and there is no regulatory concern below that level, then being slightly above an RfD that is very distant from a NOAEL on a very steep dose-response curve is not that much different from being at or below the RfD, from a decision point of view, given the uncertainty factors that generated the RfD. But being at e-prime is quite different because it is near the NOAEL. The consequences of having the effect seen at the LOAEL at either e-prime or e-doubleprime are relatively the same, given the uncertainties. So the issue becomes whether one can live with either the upper bound or the lower bound of the effect around the LOAEL.

Looking at how to combine this information with other data, the range might be sufficient to indicate whether the costs or benefits of a decision can be determined. If not, a more complex

analysis is required, unless the range is narrow enough or the effect minor enough. That is the easy decision. With the same NOAEL and LOAEL, but an RfD that is closer, there is less uncertainty and therefore the first exposure is much closer to the NOAEL. The consequences of being at the exposure e are then different. What is likely to happen and the likelihood of being wrong are different. Dr. Putzrath discussed a third case with a shallow dose-response curve, where the change of exposure does not make much difference in the response. Again, the question is whether the anticipated response is acceptable. She noted the approach is much different from the quantitative analysis she usually does, but suggested that "quick and dirty qualitative analyses" can sometimes answer the important question, which is whether the result is acceptable and cost of dealing with an effect worthwhile.

Dr. Putzrath mentioned several general issues similar to those raised by Dr. Haber. When thinking about improving uncertainty factors, it is very important to remember they are not all the same. For variability, distributional analysis has a lot of value. Distributional analyses can be done with interspecies, either chemical-specific or generic, and there are biologically based models. That might be more useful than a generic method. With LOAEL to NOAEL, there are some potential problems, and using chemical-specific data will likely give more information than a generic distributional analysis. With missing or poor quality data, additional safety factors can be used but cannot truly replace good data.

Some of the methods discussed may effectively do away with thresholds. If that is the goal, it should be explicit. There have been heated discussions about whether thresholds exist and, if they do, how to estimate them. If there is a policy choice to eliminate thresholds from the analyses, that can be done. But it should not be done by a mathematical or statistical blurring of methods.

Discussion

One participant noted there had been little consideration of risk as a function of biology and exposure. Advanced tools now exist for studying that and can give estimates of risks with less uncertainty than estimates based on default approaches. Dr. Putzrath replied that this would involve starting over, rather than improving existing methods.

Dr. Rhomberg said it is very important to use biological knowledge and advanced methods for characterizing dose responses. Distinguishing between adjustments and uncertainty about those adjustments is important to enhance the "centering" step of his approach. There remains a question about how to extrapolate from animal to human even after answering questions about variability and uncertainty.

The participant agreed that uncertainties remain even with biologically based models. But the advantage is in seeing clearly what the major sources of uncertainty are rather than leaving them hidden in the statistics. He advised care in refining a 20-year-old method.

Kenny Crump wrote an equation on a transparency to illustrate the use of lognormal distribution for extrapolating from higher to lower dose (see Appendix L). He noted that products of lognormal distributions are also lognormal. Other distributions could describe the same data with different implications at the tails, so it is important to look at the distribution assumptions very carefully.

Crump further noted the key is not risk itself but incremental risk over background. He said that this is an important distinction with important implications. He suggested the distributional approach might not be appropriate to apply to the entire population, but only to the population that would not get the disease except for the exposure. That segment of the population would likely be characterized by shallower slopes than would the entire population.

Another question is how to incorporate background. The proposed approaches start with an exposure. The "transfer function" is pharmacokinetic information that gives an internal dose. It can be represented in the model by saying, if it is greater than the threshold dose obtained from the pharmacodynamic data, there would be a response (see Appendix L). And if the threshold (internal) dose from pharmacodynamic data has a lognormal distribution, and the transfer relating external to internal dose from pharmacokinetic data has a lognormal distribution, the ratio would have a lognormal distribution. But the threshold dose cannot have a lognormal distribution because whenever there is background response, it must have mass at zero. How that mass at zero is incorporated makes a large difference in the low-dose risk result.

Dr. Crump noted that the standard log normal model (assumed by Dr. Hattis, Dr. Gaylor and others) is very flat at low dose. All the derivatives are zero at zero. But depending on how background is incorporated, it can become linear very quickly at low dose. So the argument of additivity to background should not be ruled out. He suggested that this model is just as consistent with the data as any other, so model uncertainty must be kept in mind. He recommended that any approach have a wide enough range so that linearity is included in the range that is a possible risk. He suggested looking at the large data set on PM and mortality. Those data appear linear down to the lowest observable doses. He advised taking some of these approaches, positing a level that results in a 5 percent mortality, and seeing how the approaches compare with the responses measured in the PM.

Dr. Hattis agreed with Dr. Crump that other distributions likely can be found to describe the data as well or better. Lognormal is justified not just by statistical convenience. There is also an expectation, given the many different features of people (such as chemical absorption rates and internal half-lives) that each affect susceptibility in a multiplicative way, that things will approach lognormality. The individual distributions do not have to be lognormal, they just have to interact more or less multiplicatively. Model error can never be eliminated as a concern, particularly projecting down to very low doses and risks.

Dr. Hattis said that he investigated that issue with the pharmacokinetic data. He said about 2,700 pharmacokinetic data points line up in terms of comparing a z score, suggesting they are

basically lognormal. A concentration in the upper right corner indicates a slight excess of values that would suggest a somewhat greater risk than would be projected from the lognormal in the data. Thus, for the larger data sets there could be some mixed distribution character. That means the data should be described not just with one lognormal distribution, but with two or perhaps more. That is an appropriate adjustment, rather than inventing another distribution. It makes a more complicated model, but is probably faithful to the mechanistic idea that the pharmacokinetic differences among people arise from many small factors each acting in a basically multiplicative way. Looking at all kinds of pharmacokinetic data, such as mixing volume distributions, half-lives, and so on, Dr. Hattis said his plots indicate a slight deviation from normality, which suggests slightly fatter tails at the high end than would be suggested by a lognormal distribution. He agreed that detailed modeling would be appropriate in cases of doses slightly above background.

Dr. Rhomberg agreed that additivity to background is important. In his team's methodology, they tried to distinguish where there are distributions one wants to know and specify. The team used lognormal distributions, but there are other approaches. He noted that although the same problem is found in cancer risk assessment of low doses and divergence among the tails of distributions, it will not be to the same degree.

As for Dr. Hattis' remark about lognormal being appropriate in terms of a number of factors that can be acting in a multiplicative way, Dr. Rhomberg noted the number of those factors is limited. So although the theoretical lognormal distribution goes all the way down to zero, the actual distribution from multiplying factors is limited by the number of factors. Thus somewhere in the tail of the distribution, it stops being really meaningful. Dr. Rhomberg's team chose 1/1,000 as that point where it may not longer be real.

Dr. Crump reiterated his belief that the assumption of lognormal distribution needs to be investigated. Having a finite number of factors means not having an exact lognormal distribution, but only an approximation. To get that approximation requires assumptions that are not plausible, such as that different factors in the same individual are independent. He also suggested that cost-benefit analyses would want to characterize the total risk, going down to minute exposures, which means going below the assumed lower limit of 1/1,000. There could be cases where the real impact of the analysis comes at the risk of 1/1,000,000 but where a lot of people are exposed at that level.

General Discussion

Dr. Wood read the following questions submitted by colloquium participants:

- How do we take into account multiple effects from multiple data sets for use in economic analysis?

- What is the willingness to pay question to be addressed? We are not asking what are you willing to pay so that a rat will not have a 10 percent chance of liver damage. We need this question defined for the human population.

Dr. Putzrath reiterated her question about whether thresholds are believed to exist. She said the question of low dose and how to do cost-benefit analysis in part depends on whether there is a value below which no effect is expected.

One participant said the discussions and presentations so far only addressed part of the problem. She restated Dr. Vu's colloquium objectives: to provide economists with full range of health effects associated with a chemical exposure; to define severity, onset, and duration for those effects; to identify the characteristics of the people that may be most susceptible to those effects; and to estimate the number of people at risk.

She suggested risks right at the reference dose may not be the right subject. With exposures 50 times the reference dose, the concern would be not only about a particular effect but some of the other effects seen in animal studies that were not used to derive the reference dose. Some of those issues need to be looked at. And, she said, to get at willingness to pay, economists need not just risk numbers but some advice on what kinds of effects to expect and the severity in the sensitive subpopulations.

Another discussant echoed that frustration. He suggested that Dr. Vu's objectives were the ideal data for economists, but that they might be satisfied with the same dose-specific probability for noncancer threshold effects that is available for cancer. The discussion identified more than enough tools to do that. He added that many things apply regardless of whether the effect is cancer or noncancer, including variation in the population, variations in extrapolating from animals to humans, and pharmacokinetics. But he said economists really want the slope factor for use in estimating the probability of an effect at different doses. He suggested, as an example, that economists would like to know the probability of health effects in a population living near a stack both before and after a scrubber is installed, in order to quantify the incremental impact. The question therefore becomes, why are there no dose-specific probabilities for noncancer and nonlinear threshold kinds of effects?

Multiple Endpoints

One participant replied that the problem with noncancer has nothing to do with threshold and nonlinearity or the lack of statistical methods, but is in increasing the dimensionality of the effect and in the lack of data concordance from animal species to humans. It is not a single noncancer effect, but potentially four or five effects covering a whole spectrum of target organs and levels of severity. Methods are lacking in the multivariate case. Cancer is very simple because the goal is total avoidance of any kind of cancer, ignoring the difference between lethal and nonlethal cancers. But for noncancer effects it becomes difficult to look at the risk of many different kinds of things. And there is a huge difference in costs for hospitalization for different types of

respiratory effects, for example. Pharmacokinetics provides a good handle on some aspects of differences among species, chemicals, and duration, but does not provide any risk numbers. It is still necessary to relate tissue exposure levels to toxicity, so dynamics come into play. That is where there is not enough understanding of the connection between animals and humans.

Another participant responded that pharmacokinetics helps explain the linkage between exposure and developing an adverse health effect. The discussion must include what science will provide the most accurate risk assessment, including pharmacokinetics and pharmacodynamics, to feed into an economic analysis.

Dr. Rhomberg emphasized the importance of looking at multiple effects. His team's method looks at only one effect at a time. One alternative is, in assessing impacts on a human population, to look simultaneously at all of the endpoints for which there is information, not just project from a sensitive endpoint. But there may be questions about whether each one is relevant to humans. Another problem is that the same endpoint might be measured in several different experiments, maybe in different species, with some different results. The traditional approach is to take the most sensitive, but that does not help in understanding uncertainty. One could consider how to use data showing that different doses cause liver toxicity in mice, rats, and hamsters to illuminate the distribution of differences among species. So it will be essential to look at several different endpoints, not just different degrees of severity of one endpoint. But that might require treating them as though they were separate endpoints and doing several parallel analyses in the absence of a method to model such a multidimensional response.

Dr. Vu refocused the discussion about multiple effects. She noted that the reference dose was developed to identify a dose that would cover all endpoints considered to be generally safe. Cancer is treated separately as a means of simplification. Recognizing that contaminants or mixtures could pose a myriad of effects, is it possible to look at disease outcomes, then see what dose-response information is available to predict the risk for various health endpoints, rather than lumping all noncancer into one box and looking only at critical effects? That would complicate risk assessment but would facilitate benefits analysis. If a health outcome is identified for effects thought to have biological thresholds in individuals and populations, can the available dose responses be incorporated using some of these models to give a probabilistic estimate?

One discussant suggested that if more than one effect contributes in a material way, they should be treated as separate problems, maybe correlated to some degree, but with different end effects, severities of effect, background risks, and interindividual variabilities. The aggregate effect and benefits would be the sum of those different effects.

Dr. McGartland commented that although it would be ideal to have the dose response for an RfD for noncancer effects, the draft cancer guidelines are moving toward an RfD approach for cancer, at least when the data permit a nonlinear threshold. That is a step backward to an economist because it eliminates the ability to quantify the benefits for cancer. He noted that mathematical models discussed could be applied to dose-response information at least from the RfD or

something similarly defined, like an MOE, that would, from a threshold for cancer, be a floor on which to go forward.

Another discussant emphasized that in considering multiple endpoints, the probabilistic procedures and analyses be kept very clear and as simple as possible. The benchmark dose procedures are complicated and hard for people to learn. Adding analysis for multiple endpoints will require a commitment from management of money and time, unless it is kept very simple.

Another participant pointed out that, ordinarily, dose-response information is available only for some of the endpoints of interest. The ones that are not quantified cannot be used by economists or considered by those reading their analyses. But looking at this as a multidimensional problem can provide economists a tool for incorporating the fact that certain endpoints have not been considered. It still may be necessary to look at endpoints individually as well, as is done with different cancer endpoints.

Dr. Don Barnes, Science Advisory Board (SAB) Staff Director, summarized the colloquium's progress so far. People generally have a comfortable feeling about the various methods presented. The difficulty comes in looking at the implications of their use, such as considering additivity to background and looking at multiple endpoints. In addition, economists will be challenged to determine how to value precursor effects and elicit preferences from people who do not understand what precursor effects are.

Another discussant pointed out that while the discussion has focused on cancer versus noncancer risk assessment, the Agency is trying to harmonize those approaches. From a biologist's perspective, there is no reason to develop separate methodologies for cancer and noncancer. Both include multiplicities of endpoints.

He added that the complexity is a pressing issue. Biologists are just beginning to understand the mechanism of action for cancers at the molecular level. Obtaining accurate risk assessment that reflects that biology will require complex research procedures and complicated models. To the extent that is not described accurately, there will be significant uncertainties in risk assessments and cost-benefit analyses. The complexity is embedded in the biology that underlies the risk. That is a long-range problem but one that can be approached rationally.

Dr. Hattis remarked that although his analysis was complex, the bottom line was relatively simple, in terms of power law relationships for the expected value. He said caveats on the form of the distribution are worth noting. But he suggested there is some utility in a generic analysis that can be done for appreciable numbers of chemicals and effects while the more accurate methodologies are being developed for cases with a richer database.

Another participant said the RfD/RfC-type approach has the advantage of being simple and looking at one thing. But in going above the RfD and looking at multiple endpoints, the analysis is going to change every time the exposure of interest changes, sometimes in very small

increments. With thresholds and dose-response curves, regardless of the equations and regulations, the risk will not be in a linear proportion to exposure. Even the simplest dose-response curve can get into an analysis that is complex enough for one exposure of interest. When it is 2-fold or 10-fold higher, everything could change. And that will be very complex, not just to calculate but also to communicate. One scenario might deal with liver effects, while another deals with lung cancer, and those have completely different cost structures. So it is important to think about how to do it and what to say about it.

Other Considerations

A discussant mentioned another concern with concordance of effects across species. Because most of the data are in animal species and the goal is preventing all adverse effects, the specific effect was not so critical. But trying to value a specific effect observed in animals, not knowing whether that is the response of interest in people, goes beyond any of the methodologies discussed. That issue needs to be raised in a bigger context of the type of research done for risk assessment.

Any guidance to economists should deal with exposure considerations, another discussant urged, to reflect chemicals with different routes of exposures for humans, sensitive subpopulations that may or may not be exposed, and human activity patterns that affect exposure. But one recalled that there may not be a correlation with the exposure parameters in describing distributions for RfDs or thresholds. He noted that Mr. Price found that situation in a risk assessment he did of PCBs where he distributed the human interindividual adjustment factor.

One participant raised an additional caveat, a concern about how well sensitive populations have been studied. The distributions can become bumpy and speculative and many of the data are from small data sets. The issue of background is also really important and needs to be tackled in general in trying to characterize nonlinear effects, not just for benefits analysis.

Dr. Vu summarized the discussion, noting that the next day's meeting would focus on developing an agenda for how to address these issues, identifying available tools and methods, and getting economists and risk assessors to work together to address some of these issues. She asked the group for suggestions. Although research into biological risk assessment models is further enhancing the understanding of the mechanism, EPA must use the standard toxicological information that is available. Approaches and methods can be developed to address some of the questions of economists, and can also improve the ability to characterize risk to inform risk management decisions.

Suggestions for Moving Ahead

One discussant suggested doing a case study or a series of case studies. Another suggested following the Air office's example of starting with the economists' need and working backwards to see how many human data are really available, using clinical studies, hospital studies, epi

studies, and occupational health investigations. Identifying different health endpoints, and finding environmental exposures that might be linked to a health effect, can identify what information is required to give economists what they need. It might also identify the kind of animal studies needed to supplement that information base.

Dr. Rhomberg reminded the group of the saying, "the perfect is the enemy of the good." Doing it exactly right may not get to the goal, so the question becomes, what is practical to try to do? He advised working within a structure that, to the extent possible, does not bind you to a particular choice that is made now for expediency. Begin to think about multiple endpoints, but consider how often that would be a critical thing to worry about. He noted that his team's sensitivity analysis found some assumptions were really important to try to improve, whereas others were too small to be a factor affecting satisfaction with the approach. That kind of sensitivity analysis can help prioritize.

Dr. Crump suggested that a study of model uncertainty could be done fairly simply using existing data to identify what models are consistent with the data and what risks they would indicate. That would be similar to sensitivity analysis but with a slightly different focus.

Dr. Haber mentioned precursors and biomarkers, hot areas of research that can be used to extrapolate the dose-response curve to lower dose levels and to give information about what is going on at environmentally relevant exposure levels. Those precursor effects may be measurable in human populations that are exposed. If they can be measured quantitatively, this can assist in quantifying the real risk.

One discussant pointed out that, if the goal is to do better risk assessments for use by the economists, it is important to know how much accuracy they need. That is different from doing better risk assessments for the purposes of better risk assessments. A crude range that encompasses a fair amount of sensitivity, identifying which parameters are important, might be much more useful than better refinements of distributions or other parameters of models, especially when branching out to multiple endpoints. One type of risk analysis can get in the range between 1/1,000 and 1/10,000; an entirely different type is needed if the difference between 1/10,000 and 2/10,000 matters to the economist. Where refinements are made depends on what they need.

Another suggested the idea of giving economists a worst case so they can determine whether there is any economic impact, or a best case to begin looking at an overall range, rather than trying to begin with perfect knowledge.

Another discussant noted that whatever methodology EPA chooses will have an impact on all the State EPA offices. She suggested that incremental changes are easier to adapt to, particularly given the growing complexity of risk assessment methodologies and the need to communicate the answers to the public.

One participant suggested that there might be a benefit in focusing more on the qualitative descriptive aspects of the chemicals, rather than just on the quantitative ones. That might help economists in the risk characterization portion. She inquired whether there are certain categories of outcome with which they are particularly concerned. It is possible to describe qualitatively the likelihood that a particular chemical exposure would cause a certain type of outcome.

She added that there is a lot of room to improve understanding of biology, biomarkers of exposure, and precursors. But very carefully done studies in humans are needed before they can be used in a truly predictive way. This may be the time to really start assessing how and where the money goes for the studies. It is possible to leverage that investment, such as doing exposure assessment work in the context of an ongoing study.

Dr. Hattis supported that view, saying that his data were not designed for the purpose for which he used them. But to really know about all of the interactions with different subpopulations according to age, gender, and illness categories, it would be helpful to actually collect new data deliberately designed with stratified random samples that can measure some of these parameters and response functions in a more deliberate way.

One participant supported the idea of using biological and chemical-specific data wherever possible in evaluating variability and uncertainty, but added that the level of detail is not always there to do biological modeling. So in developing methods, it is important to look at how to treat the large uncertainties of an incomplete database.

Another said that, having looked at all the issues, the goal is still attainable. Case studies might help, as could a chart of information for the economists that includes factors such as the major effects of concern; information on mode of action; some kind of likelihood that the mode of action might be active in humans; what the expected contact rate would be with segments of the population, as well as a sense of the toxicity; what the concentrations might be in a medium and what kind of contact with that medium would be expected. With that chart and the risk numbers, economists could determine how many cases of some kind of effect would be seen. That package would have to include a translation into human health conditions, given an animal effect and what is known about mode of action, pharmacokinetics, even chemical structure and effects for similar chemicals. That could be made into guidance on whether the effects would be expected in humans, what is the real exposure, and what is the real likelihood of toxicity in a human. It is not as simple as providing a slope factor that can be applied to the whole population of the United States. That is how the really large numbers come about.

Adjournment

Dr. McGartland noted that he remains optimistic. There are some models with which to begin making headway, at least for single effects, toward meaningful cost-benefit analysis. It is important to put these benefits in context, even if they cannot be valued, and cases or symptoms mean a lot more than some of the other measures. He thanked participants from outside EPA for

their presentations and comments. Over the long run, he said, he could see lobbying for more joint research on the issues.

In the short run, Dr. McGartland said that he was intrigued with the case studies approach, perhaps starting with single effects and going on from there, and perhaps even coming up with some guidelines for standardizing the approach, with appropriate caveats about model selection, and so on, and what they mean for the benefit estimates. But to do that, a lot more work is required. One or two case studies will not be sufficient to provide the necessary sense of robustness and comfort with the cost-benefit analyses and risk assessments that go into the public domain. The ideal for economists would be a continuous curve that allows for talking about margins. But they can work with far less.

His hope for day two is to talk about more concrete steps in both the short run and the longer run. He said there would be an opportunity for colloquium participants to join in further work. He concluded by noting that this was the first meeting where economists and risk assessors have tried to solve a common problem. He called the chart idea intriguing and a good way to move forward and help economists articulate the kinds of information risk assessors can provide. He said the work will pay dividends quickly.

Dr. Wood closed the meeting by thanking the participants.

Appendix A

External Participants

Colloquium on Approaches to Quantifying Health Risks for Threshold or Nonlinear Effects at Low Dose
September 28, 2000

External Participants

Sandra Baird
The Baird Group
36 Duffield Road
Auburndale, MA 02466-1004
(617) 527-9868 (v)
(617) 527 4235 (f)
sbaird@world.std.com

Kenny Crump
ICF Consulting
602 East Georgia Avenue
Ruston, LA 71270
(318) 242 5019 (v)
(318) 255 4960 (f)
kcrump@icfconsulting.com

Dave Gaylor
Sciences International
13815 Abinger Court
Little Rock, Arkansas 72212
(501) 228-9773 (v)
(501) 228-7010 (f)
dgaylor@sciences.com

Lynne Haber
Toxicology Excellence for
Risk Assessment (TERA)
1757 Chase Avenue
Cincinnati, OH 45223
(513) 542 7475 Ext. 17 (v)
(513) 542-7487 (f)
haber@tera.org

Dale Hattis
Center for Technology, Environment and
Development
Clark University
950 Main Street
Worcester, MA 01610
(508) 751-4603 (v)
(508) 751-4600 (f)
dhattis@aol.com

Paul Price
Ogden Environmental and Energy Services
15 Franklin Street
Portland, ME 04107
(207) 879-4222 (v)
(207) 879-4223 (f)
psprice@oees.com

Reisha Putzrath
Georgetown Risk Group
3223 N Street N.W.
Washington D.C. 20007
(202) 337-8103 (v)
(202) 342-2110 (f)
rmputzrath@mindspring.com

Lorenz Rhomberg
Gradient Corporation
238 Main Street
Cambridge, MA 02142
(617) 395-5000 (v)
(617) 395-5001 (f)
lrhomberg@gradientcorp.com

Appendix B

Participant List

Colloquium on Approaches to Quantifying Health Risks for Threshold or Nonlinear Effects at Low Doses
September 28, 2000

Participant List

Rebecca Allen
Jihad Alsadek
Dan Axelrad
Sandra Baird
Don Barnes
Steven Bayard
Nancy Beck
Robert Beliles
Dave Bennett
John Bennett
Lynne Blake-Hedges
Tracey Bone
Ethel Brandt
Marilyn Brower
Susan Carillo
David Chen
Jim Cogliano
Gary Cole
Rory Conolly
Marion Coply
Kenny Crump
Linda Cullen
Vicki Dellarco
Chris Dockins
Julie Du
Gary Foureman
Dave Gaylor
Jeff Gift
Lynne Haber
Trish Hall
Dale Hattis
Rick Hertzberg

Richard Hill
Lee Hoffman
Jennifer Janoit
Barnes Johnson
Mark Johnson
Jin Kim
Gary Kimmel
Steve Knott
Arnie Kuzmak
Elizabeth Margosches
Alec McBride
Al McGartland
Robert McGaughy
Patricia Murphy
Deirdre Murphy
Onyemaechi Nweke
Ed Ohanian
Marian Olsen
Dan Olson
Nicole Owens
Fred Parham
Resha Putzrath
Lorenz Rhomberg
Paul Rice
Rita Schoeny
Jean Schuman
Jennifer Seed
R. Woodrow Setzer
Nathalie Simon
Ted Simon
Judy Strickland
Linda Teuschler
Vanessa Vu
Pauline Wagner
Ann Watkins
David Widawsky
Diane Wong

Bill Wood
Tracey Woodruff

Appendix C

Agenda

United States
Environmental Protection Agency
Risk Assessment Forum

Colloquium on Approaches to Quantifying Health Risks for Threshold or Nonlinear Effects at Low Dose

Omni Shoreham Hotel
2500 Calvert Street N.W.
Washington D.C. 20004

September 28, 2000

Agenda

Colloquium Co-Chairs: **Al McGartland and Vanessa Vu**

8:30AM	**Registration**
9:00AM	**Welcome** *Bill Wood, EPA Risk Assessment Forum*
9:05AM	**Perspectives: A Risk Assessor's Point of View** *Vanessa Vu*, National Center for Environmental Assessment
9:25AM	**Perspectives: An Economist's Point of View** *Al* McGartland, National Center for Environmental Economics
9:45AM	**Dose-Response Based Distributional Analysis of Threshold Effects** *Lorenz Rhomberg, Gradient Corporation* *Sandra Baird, The Baird Group*
10:15AM	**Characterizing Risks Above the Reference Dose** *Paul Price, Ogden Environmental and Energy Services*
10:35AM	**Expected Values of Population Dose Response Relationships Inferred from Data on Human Interindividual Variability in PK and PD Parameters** *Dale Hattis, Clark University*
10:55AM	BREAK
11:10AM	**Interindividual Sensitivity** *Dave Gaylor, Sciences International*
11:30AM	**Use of the Categorical Regression Methodology to Characterize the Risk Above the RfD** *Lynne Haber, Toxicology Excellence for Risk Assessment (TERA)*

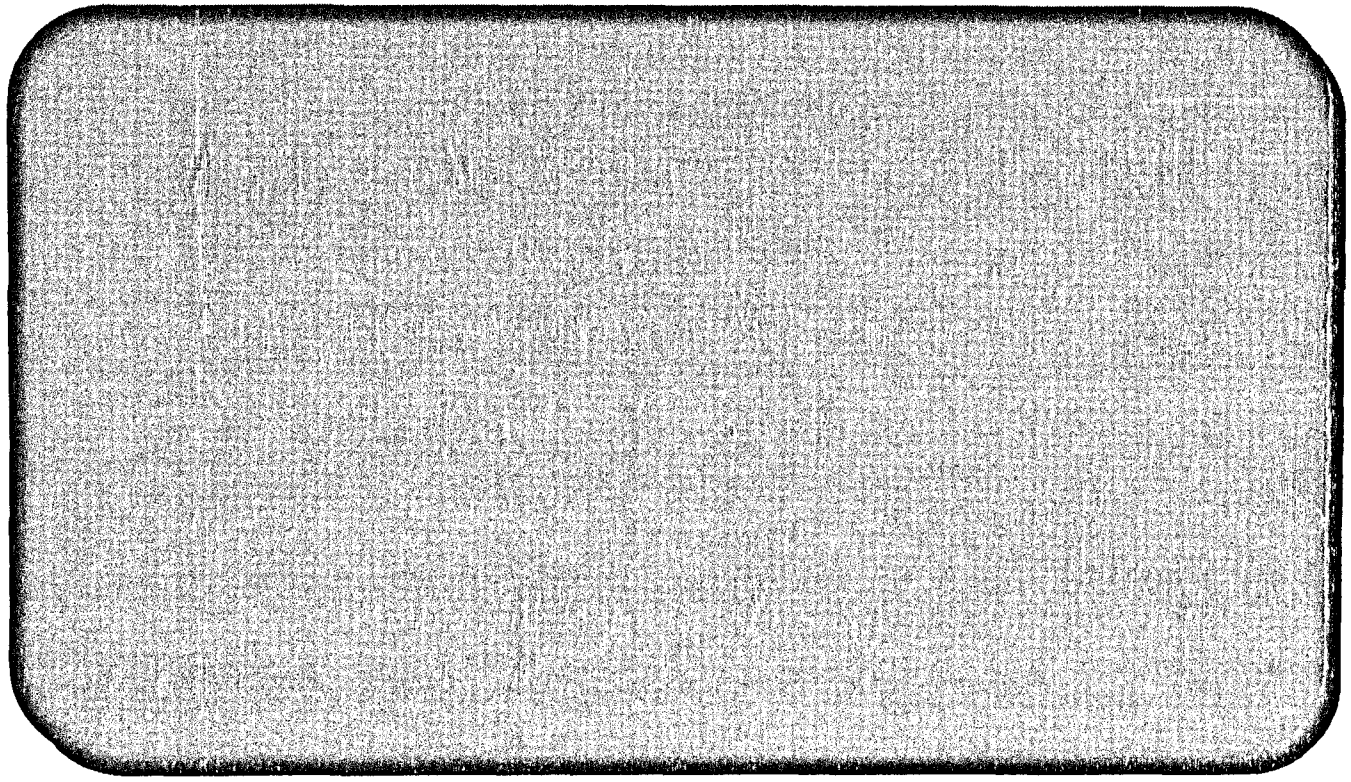
11:50AM	**Risks Between the LOAEL and the RfD/RfC: A Minimalist's Approach** *Resha Putzrath*, Georgetown Risk Group
12:10PM	L U N C H (on your own)
1:15PM	**Facilitated Roundtable Discussion** *Moderator: Bill Wood*
	(BREAK 3:00 - 3:15PM)
4:30PM	**Concluding Comments and Next Steps** *Vanessa Vu and Al McGartland*
5:00PM	A D J O U R N

Appendix D

Presentation Overheads

Vanessa T. Vu
National Center for Environmental Assessment
Office of Research and Development
U.S. Environmental Protection Agency

Risk Assessment Forum
National Center for Environmental Economics
September 28, 2000



Vanessa T. Vu, Ph.D.
National Center for Environmental Assessment
Office of Research and Development
U.S. Environmental Protection Agency

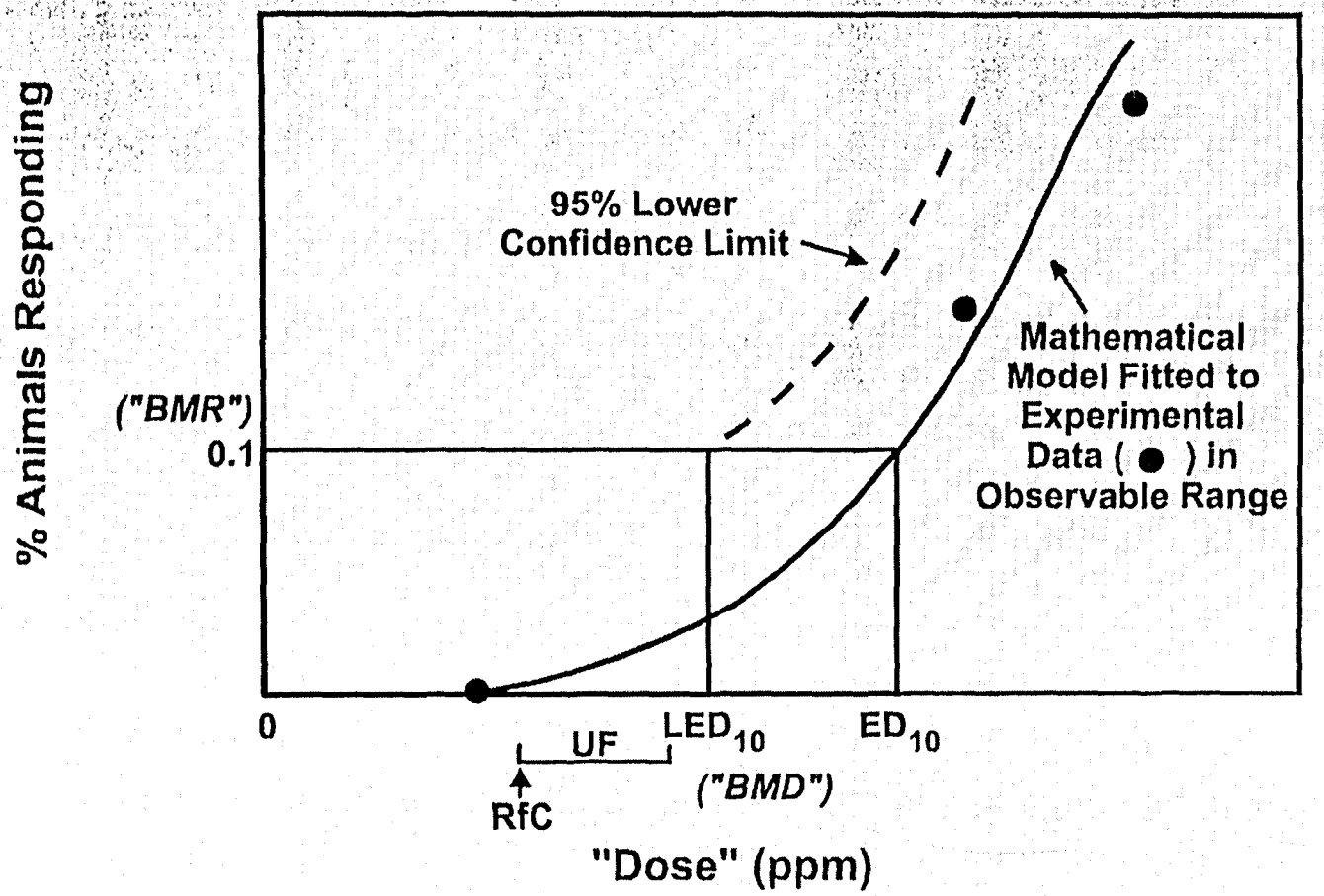
Outline

- Background
  - Issues in human health risk assessment & valuation of health benefits
- Objectives and Structure of Colloquium

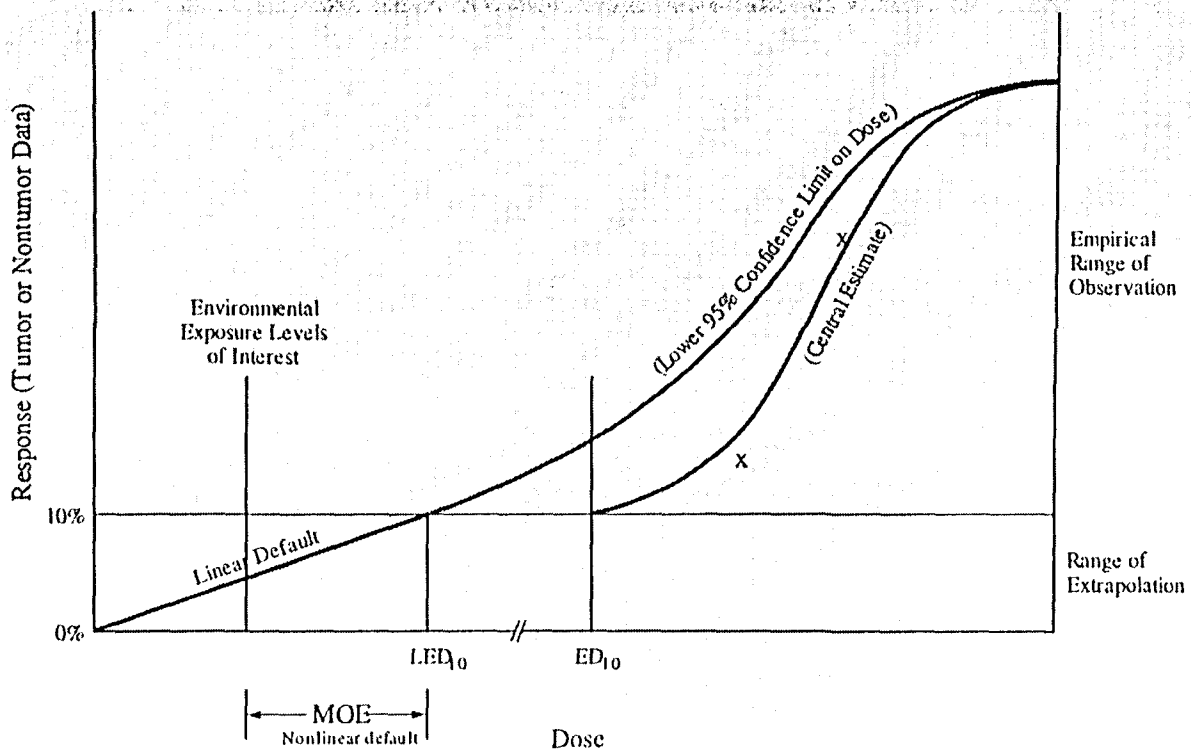
Current Efforts in Improving Risk Assessment

- Harmonized and integrated approaches for all health endpoints
  - Emphasis of mode of action (MOA)
  - Two-step dose response assessment
  - Chronic and less-than lifetime exposures
- Fuller characterization of risks
  - Probabilistic estimates, RfD/C, margin of exposure (MOE)
  - Susceptible populations

"Benchmark Dose" Approach to Dose Response Analysis for Noncancer Endpoints



Dose Response Assessment



Health Benefit Analysis

Information Needs

- Characterization of a full range of health effects potentially associated with a contaminant(s)
- Nature of specific effects- *e.g. severity, onset, duration*
- Characteristics of people potentially affected- *e.g. age, health status*
- Estimation of number of people at risk

Issues Related to Valuation of Health Benefit Analysis

- *Near-term*
  - RfD/C & MOE methods do not provide quantitative estimates of risk below POD
  - RfD/C focuses only on critical effect from chronic exposure
  - Critical effects need to be related to adverse human health outcomes
  - MOE for cancer based on precursor effects

Issues Related to Valuation of Health Benefit Analysis

- *Emerging Issues*
  - Increased use of biomarkers of effect and susceptibility
  - Valuation of more subtle effects

Colloquium Objectives

- Explore possible approaches for quantifying risks below POD
  - Biological thresholds
  - Nonlinear dose-response curve at low dose

Colloquium Structure

- Overview of approaches to health benefit analysis
- Presentations of available approaches and methods to quantify risks below POD
- Roundtable discussion
  - Identify tools and methods for near-term use

Appendix E

Presentation Overheads

Al McGartland
National Center for Environmental Economics
Office of Policy, Economics and Innovation
U.S. Environmental Protection Agency

Benefits Analysis at EPA

Al McGartland

Director

National Center for Environmental Economics

Office of Policy, Economics and Innovation

Outline

- Economic Analysis Guidelines
- Background: Benefit-Cost Analysis
- Benefits Analysis Methods
- Criteria Air Pollutant Examples

Guidelines for Preparing Economic Analyses

- EPA's Science Advisory Board, comprising leading environmental economists from major universities and research institutions, reviewed the Guidelines throughout its development for accuracy in both economic theory and practice.
- In their final report, the Board gave the Guidelines an overall rating of "excellent," saying, they "succeed in reflecting methods and practices that enjoy widespread acceptance in the environmental economics profession."

Growing Demand for Economic Analysis at EPA

- Executive Orders and legislation increasingly require economic analysis of Agency rules:
  - E.O. 12866
  - Thompson Language
  - Safe Drinking Water Act
  - SBREFA
  - UMRA
  - Proposed Regulatory Reform bills

Multiple Decision Criteria Consistent with Economics

- Factors in Decision Making Process
  - Ethics
    - Distributive Justice
    - Environmental Justice
  - Sustainability
  - Political Concerns
  - Legal Consistency
  - Institutional Feasibility
  - Technical Feasibility
  - Enforceability
  - Efficiency (Benefits/Costs)

Why use Benefit-Cost Analysis?

- Benefit-Cost Analysis attempts to simulate a private market test on the production of public goods (e.g., environmental protection)

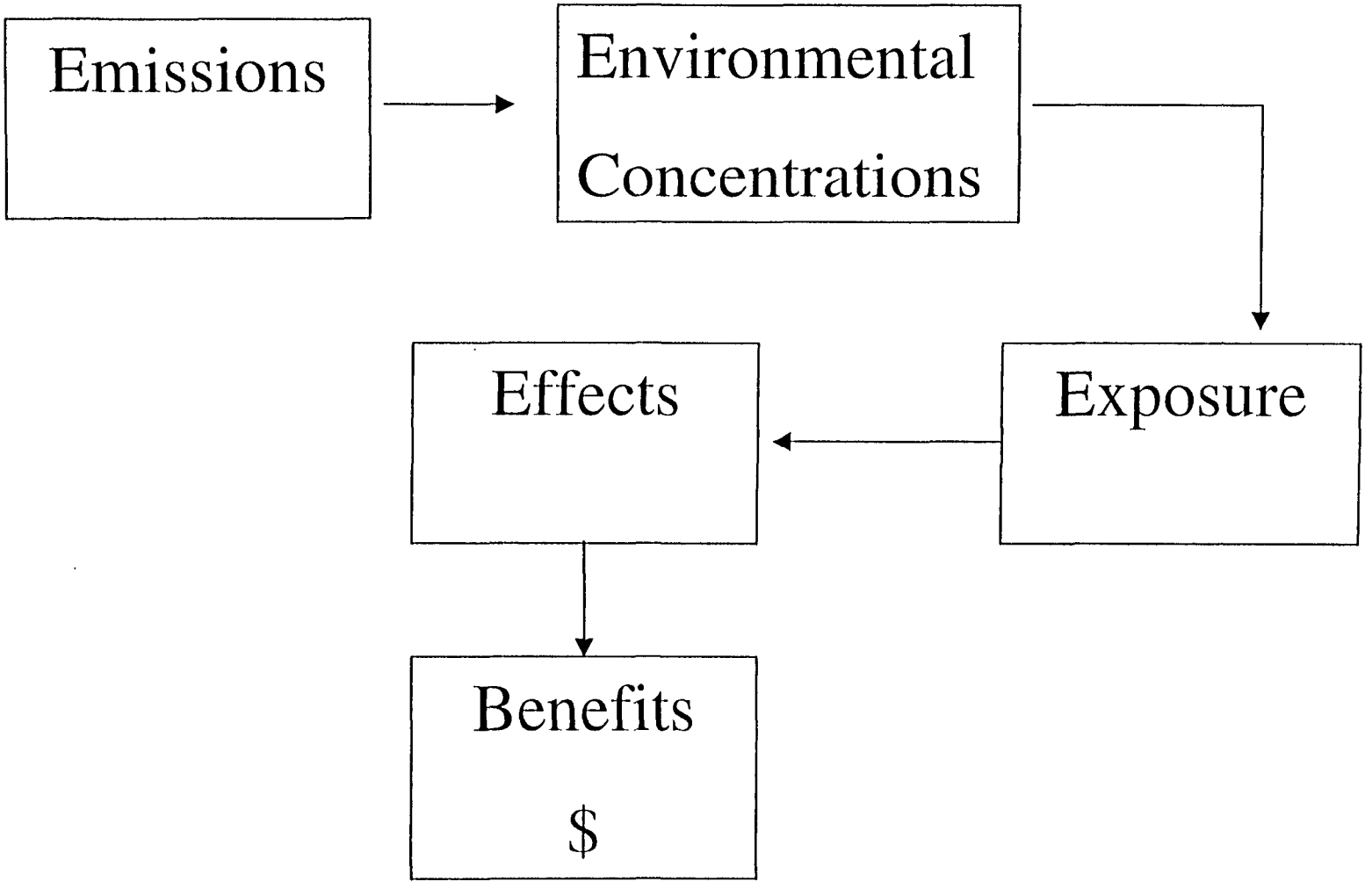
Private Market Test

- Private Markets allocate resources to efficient uses automatically
- If a manufacturer cannot sell its output for more than it costs to produce, it goes out of business.
  - This manufacturer was an "inefficient" user of society's scarce resources.
  - The cost of the resources used in production was greater than the value of goods produced.
  - Discipline of the private market forces this inefficiency out of the system and rewards efficient users of resources - i.e., those who create net positive value or net benefits.

Approaches to Benefits Analysis

- Damage Function Approach: estimate reduced incidence of adverse effects, multiply by estimated value per case avoided
- Indirect Approach: value an environmental improvement understood in general terms (e.g., Exxon Valdez damages assessment)

Steps in Benefits Analysis



Implementing the Damage Function Approach: Valuation Methods

- Cost-of-Illness
- Averting Behaviors
- Hedonics (e.g., wage-risk tradeoffs)
- Stated Preference (survey methods)

Alternative Benefits Methods

- Potential approaches to valuing human health effects without using damage function approach:
  - Value exceedances of the RfD/RfC
  - Value "peace of mind" or hypothetical insurance policy
- Potential problem with these approaches: how to express the values in terms of marginal changes

Valuation Examples--Criteria Pollutants

- The benefits estimates for criteria air pollutants are the most extensive that EPA has produced
- Dose-response functions for many effects
  - able to quantify reduced incidence of: mortality, bronchitis, asthma, hospital admissions, respiratory symptoms, etc.
- Conducted under Section 812 of the CAA--which mandated extensive SAB review

Health Benefits from Air Pollution Control (Selected Effects)

Health Effects	Annual Cases Avoided (2010)
Premature Mortality	23,000
Chronic Bronchitis	20,000
Hospitalizations	
-Respiratory	22,000
-Cardiovascular	42,000

Benefits--Unit Values

Health Effect	Mean Value per Case Avoided (1990 $)	Type of Valuation Study
Mortality	$4,800,000	Wage-risk, stated preference
Chronic Bronchitis	$ 260,000	Stated preference
Hospitalization		
-Respiratory	$ 6,900	Cost of Illness
-Cardiovascular	$ 9,500	Cost of Illness

Results of Human Health Benefits Valuation, 2010

Health Effects	Monetary Benefits (in millions 1990 $)
Premature Mortality	$100,000
Chronic Bronchitis	$ 5,600
Hospitalizations	
-Respiratory	$ 130
-Cardiovascular	$ 390
Other Health Effects	$ 2,000

Conclusions

- Demands for economic analysis at EPA are increasing
- Dose-response functions are a critical input to quantification of benefits
  - Alternate benefits methods a possibility
- Effects without dose-response functions are unquantified in benefit-cost analysis, and thus perceived by some as "not counting"

Appendix F

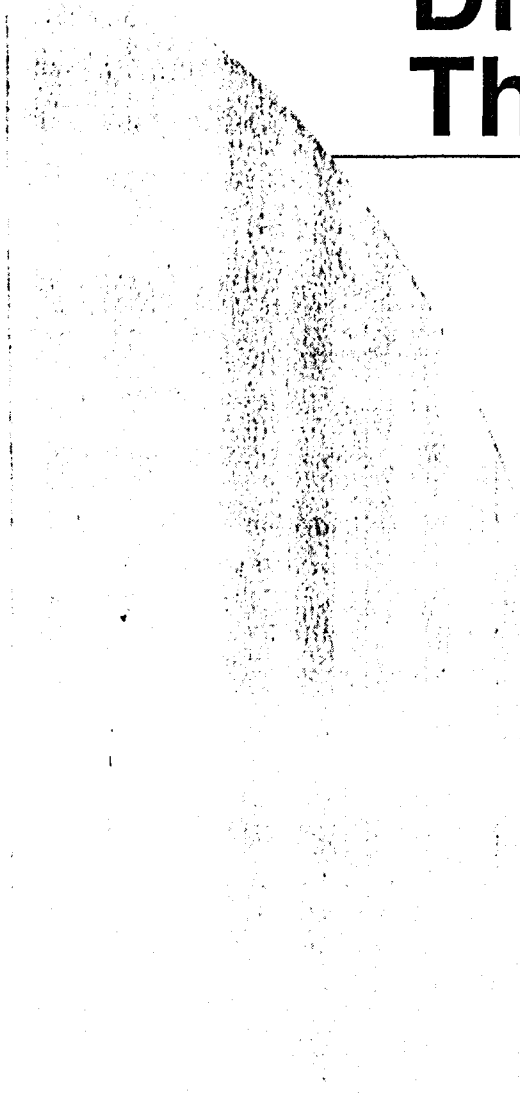
Presentation Overheads

Sandra Baird
The Baird Group

and

Lorenz Rhomberg
Gradient Corporation

Dose-Response Based Distributional Analysis of Threshold Effects



Sandra J.S. Baird
Lorenz Rhomberg

Collaborators:

John S. Evans, Paige Williams, Andrew Wilson

Overview

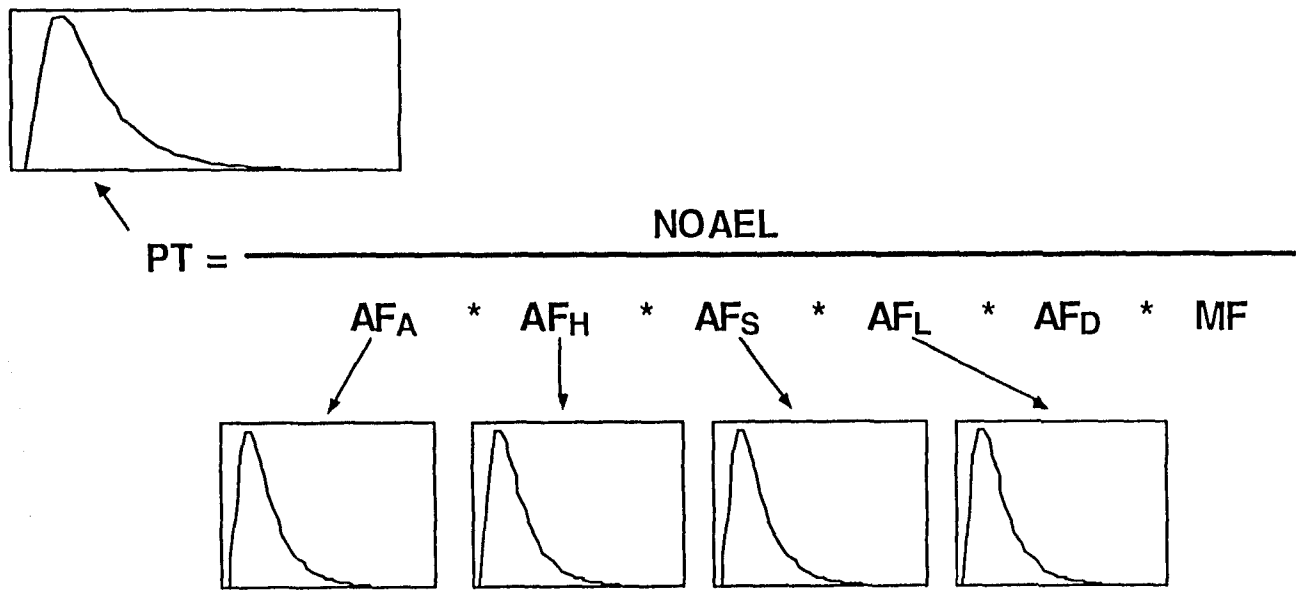
- Why move from existing approach
- Mental models
- Framework for distributional approach
- Underlying theoretical and empirical support
- Case study: Ethylene Oxide
- Benefits of Distributional D-R approach

Why Change?

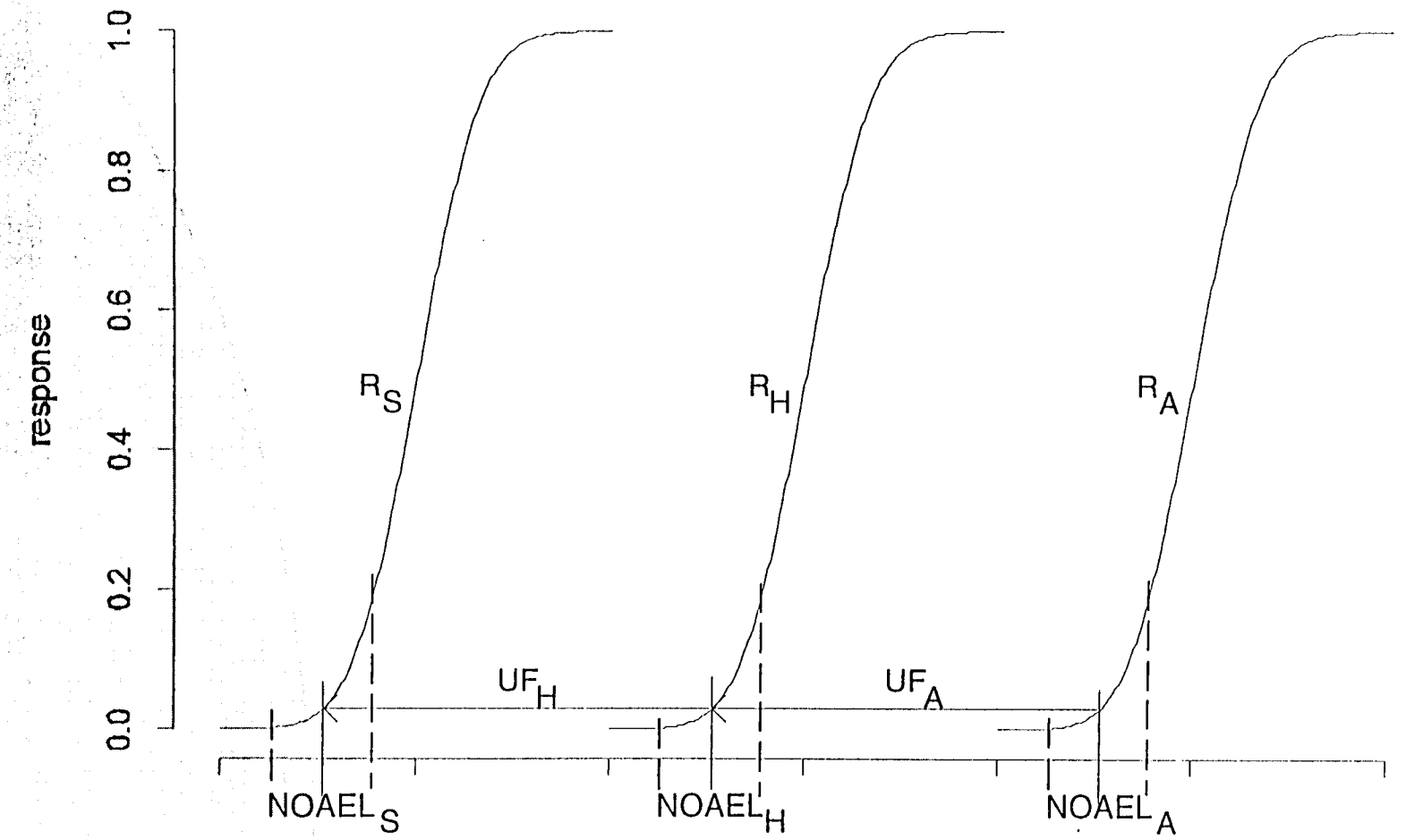
- Uncertainty in RfD is unknown
- Protection at RfD is unknown and inconsistent
- Risk assessment and risk management are intertwined
- Risk from exposures greater than RfD is unknown
- Prevents informed decision making

Current Mental Model

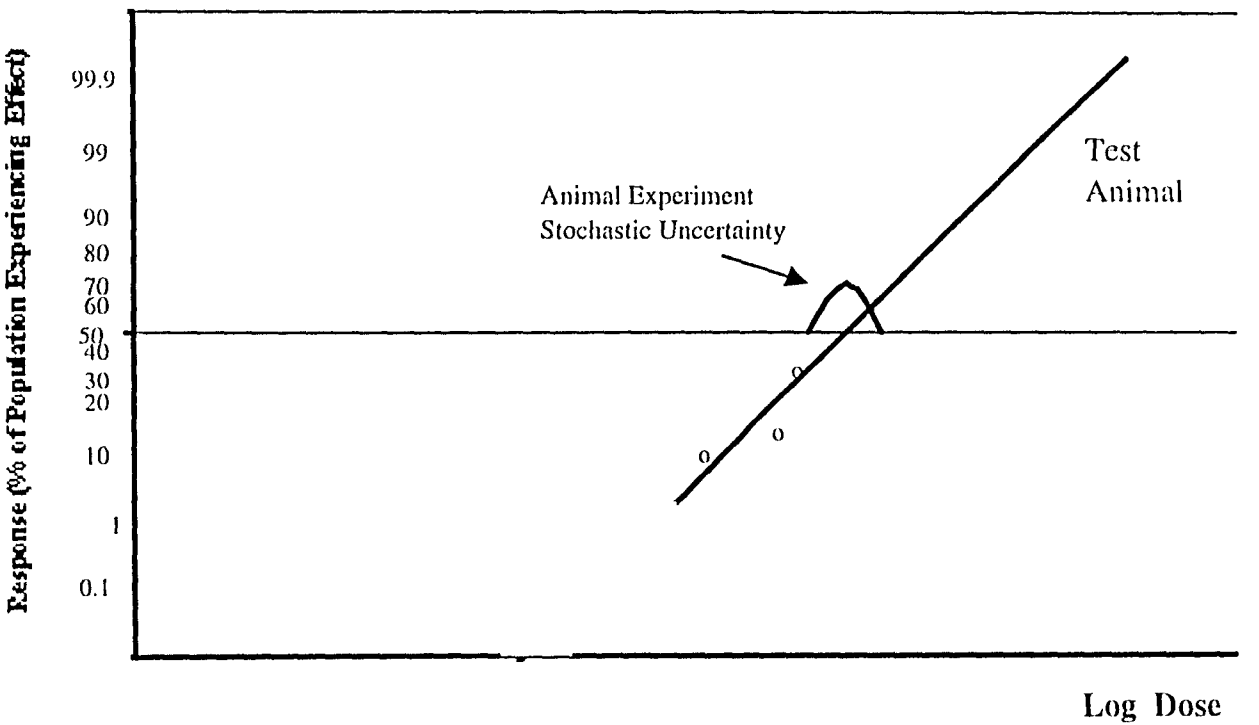
$$RfD = \frac{NOAEL}{UF_A * UF_H * UF_S * UF_L * MF * D}$$



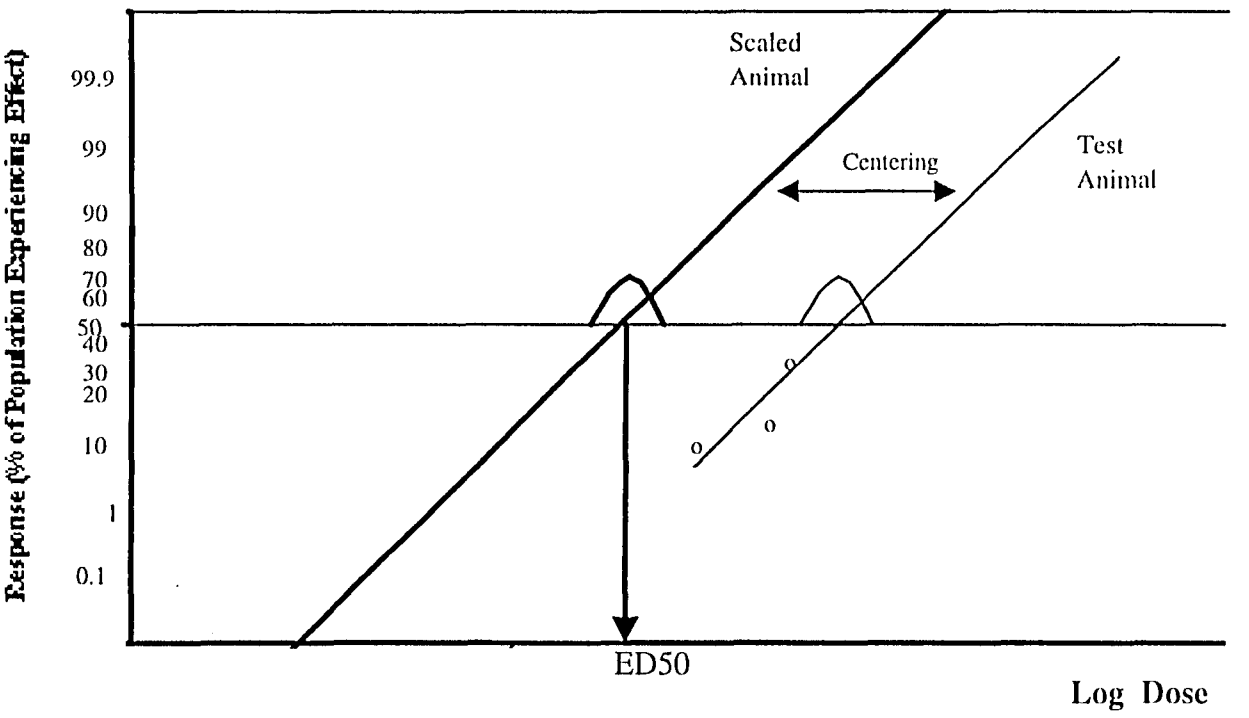
Current Mental Model



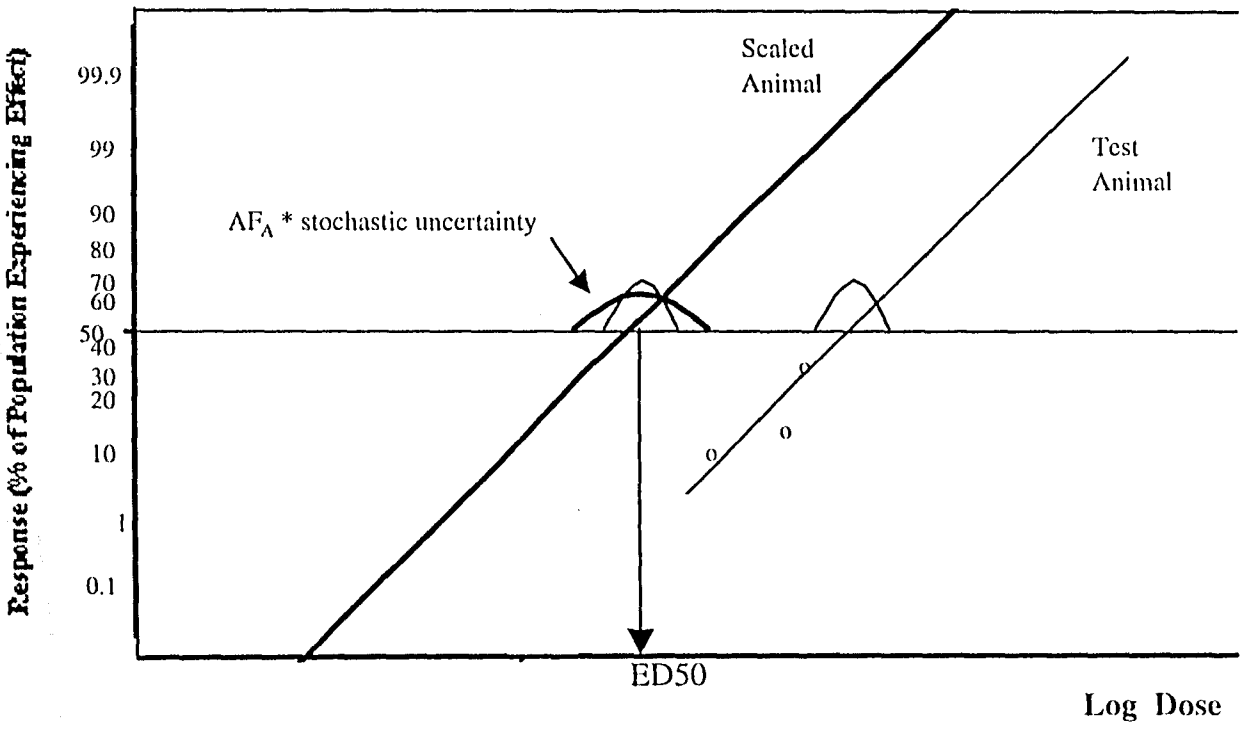
Dose-Response Based Mental Model



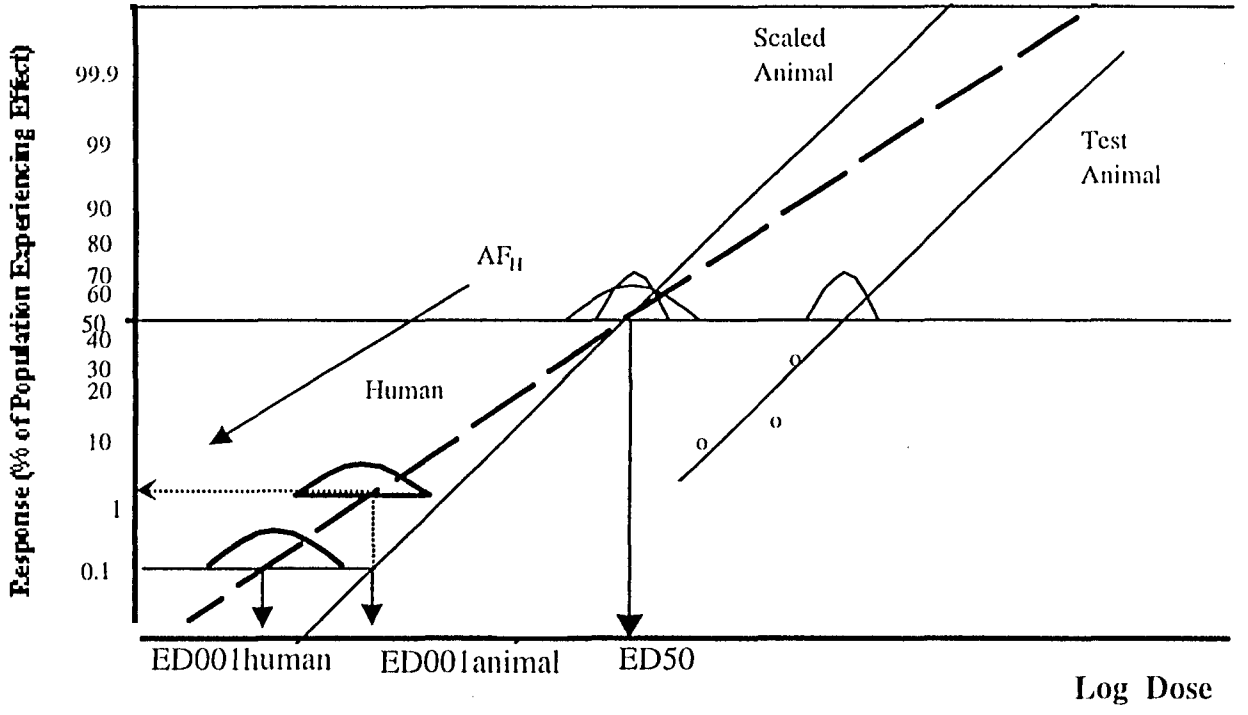
Dose-Response Based Mental Model



Dose-Response Based Mental Model



Dose-Response Based Mental Model



Framework for Distributional Approach

$$E\tilde{D}_{xh} = \frac{E\tilde{D}_{xa}}{A\tilde{F}_A}$$

$$RfD \approx E\tilde{D}_{yh} = \frac{E\tilde{D}_{xh}}{A\tilde{F}_H}$$

RfD is determined by choosing level of population risk, *y*, and by choosing level of confidence from the uncertainty distribution of the dose estimated to yield *y* response.

Theoretical and Empirical Support

- Animal to Human - AF_A
  - Scaling ("Centering")
  - Estimates of uncertainty
- Human Heterogeneity - AF_H
  - Probit slope
  - Estimates of human population variability

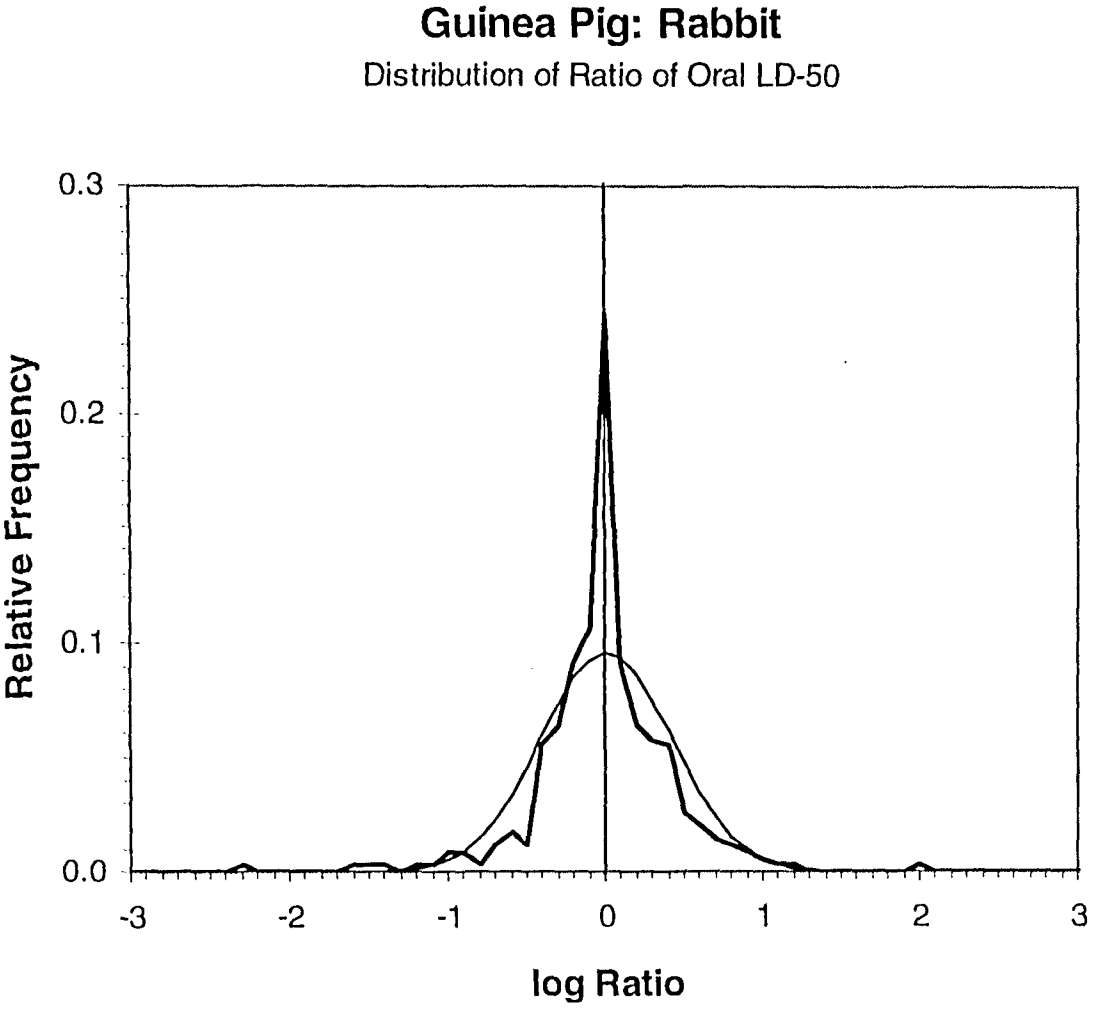
Dose Metrics: Scaling

- Allometric - BW^b
- RfC - HEC
- Chemical specific - PBPK

Estimates of Uncertainty In Scaling

Source	Reference	Range of GSDs
Pesticide NOAELs	Baird et al., 1996	4.1 - 4.9
LD50s	Rhomberg & Wolff, 1999	2.5 - 6
Antineoplastic Agent MTDs	Schmidt et al., 1997	2.6 - 3.7

Uncertainty in Scaling



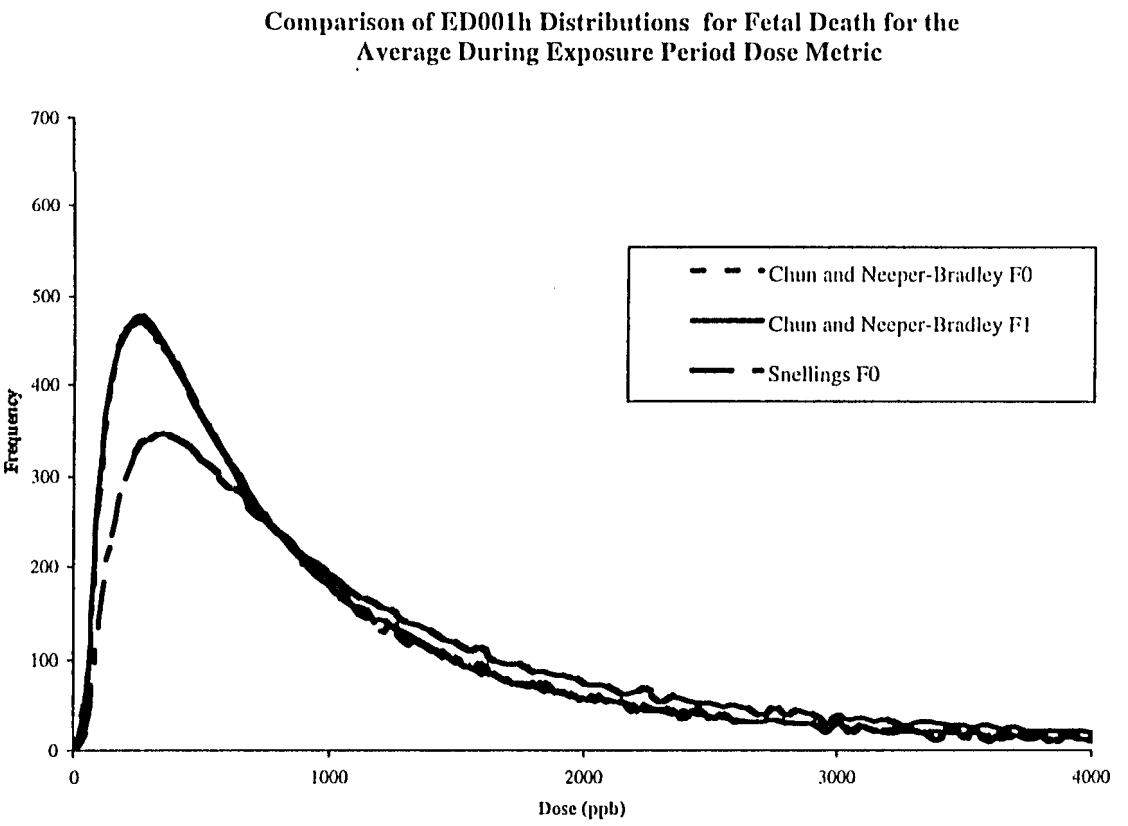
Human Heterogeneity

- Based on tolerance distribution
- Assumes humans are more variable in response thresholds than test animals
- Variability described by log-probit slope of dose-response curve (GSD)
- Slopes derived from human data (Hattis et al., 1999)

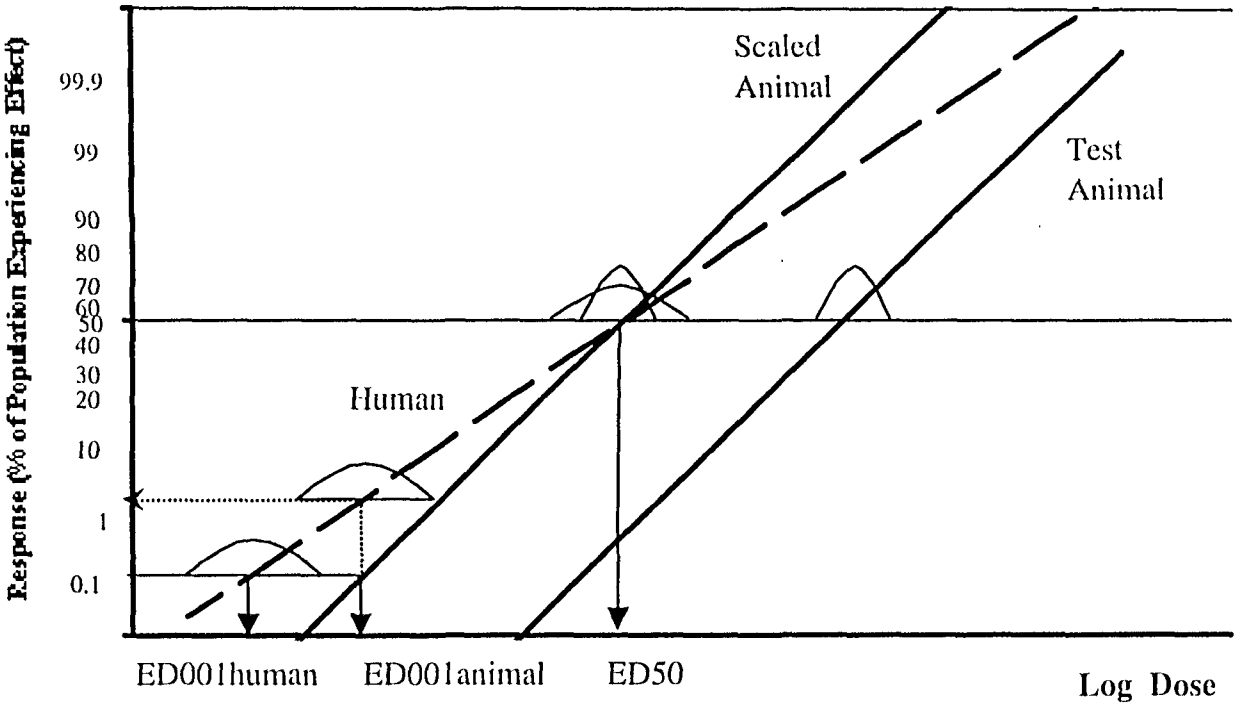
Case Study: Ethylene Oxide

- Studies: 2 Developmental
  2 Reproduction
- Endpoints: post implantation loss, fetal/pup body weight
- Advanced statistical d-r models
- Empirical estimates of uncertainty
- Sensitivity analysis

Case Study: Ethylene Oxide



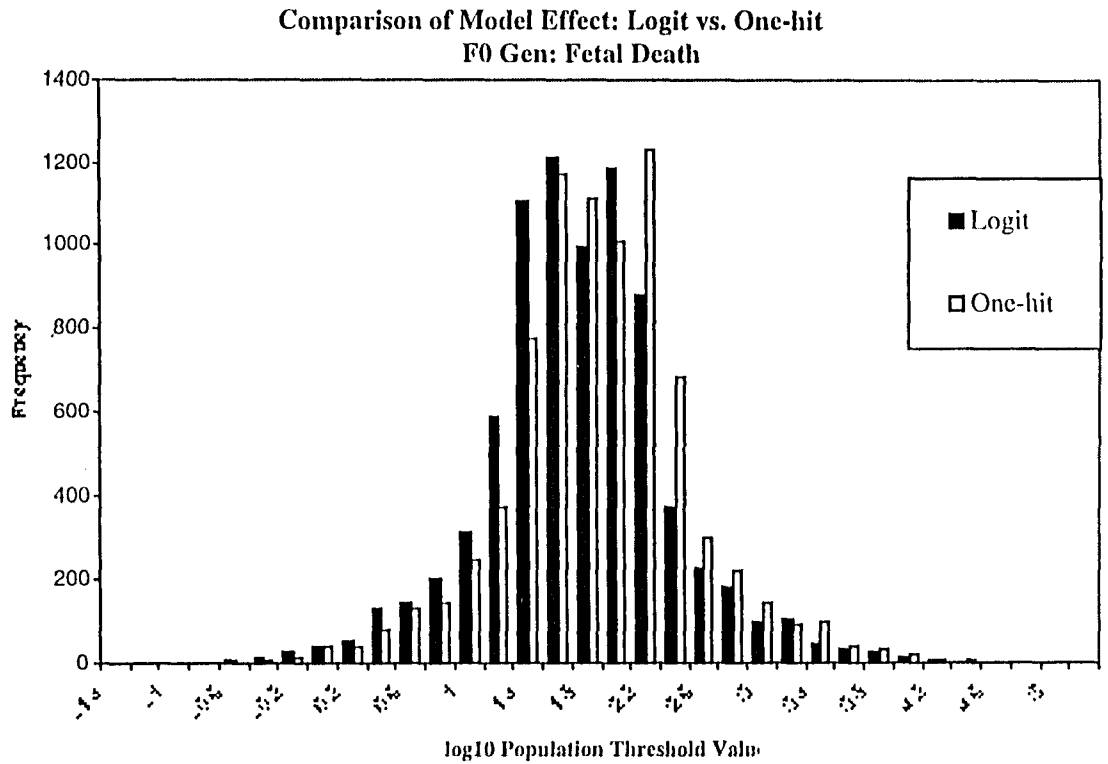
Dose-Response Based Mental Model



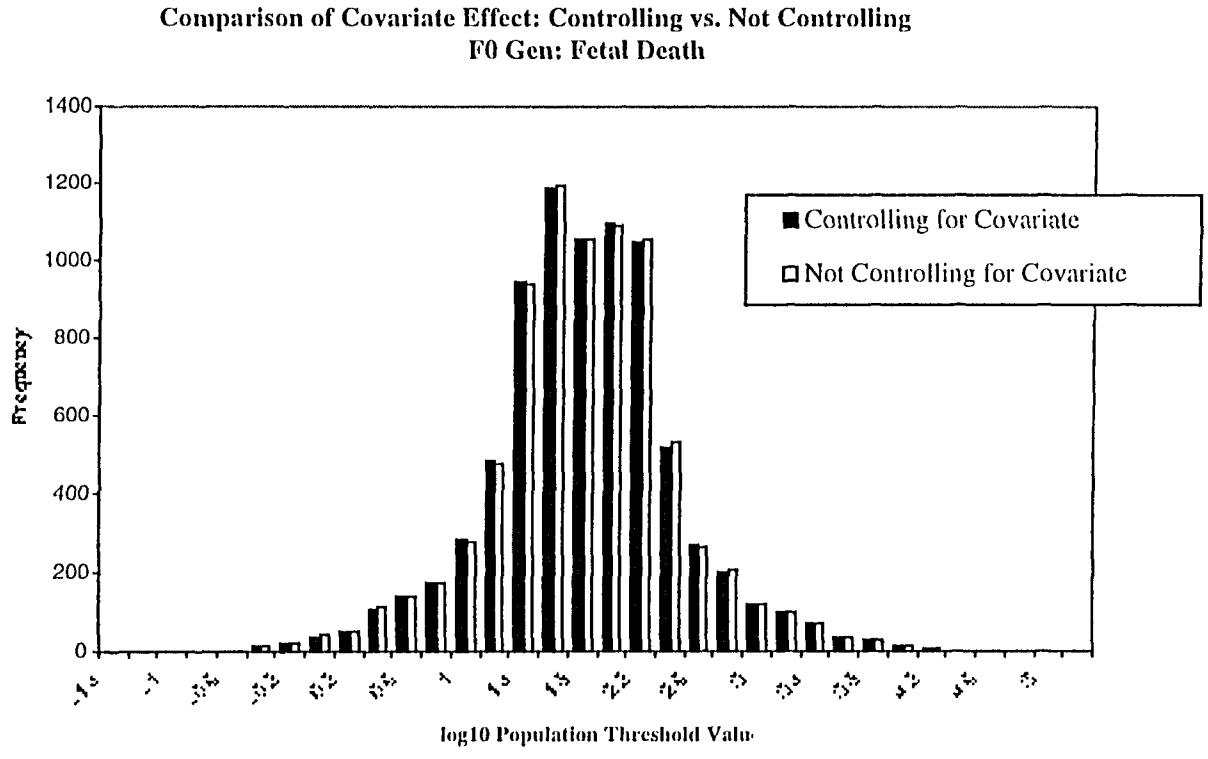
Ethylene Oxide: Results Summary

Endpoint	RfC LED10a/30 (ppb)	ED10a (ppb)	LED10a (ppb)	ED001h (ppb)	LED001h (ppb)
5% Body Wt. Reduction (Devel.)	180	7,000	5,400	290	46
Fetal Death (Repro.)	200	6,800	6,000	700	120

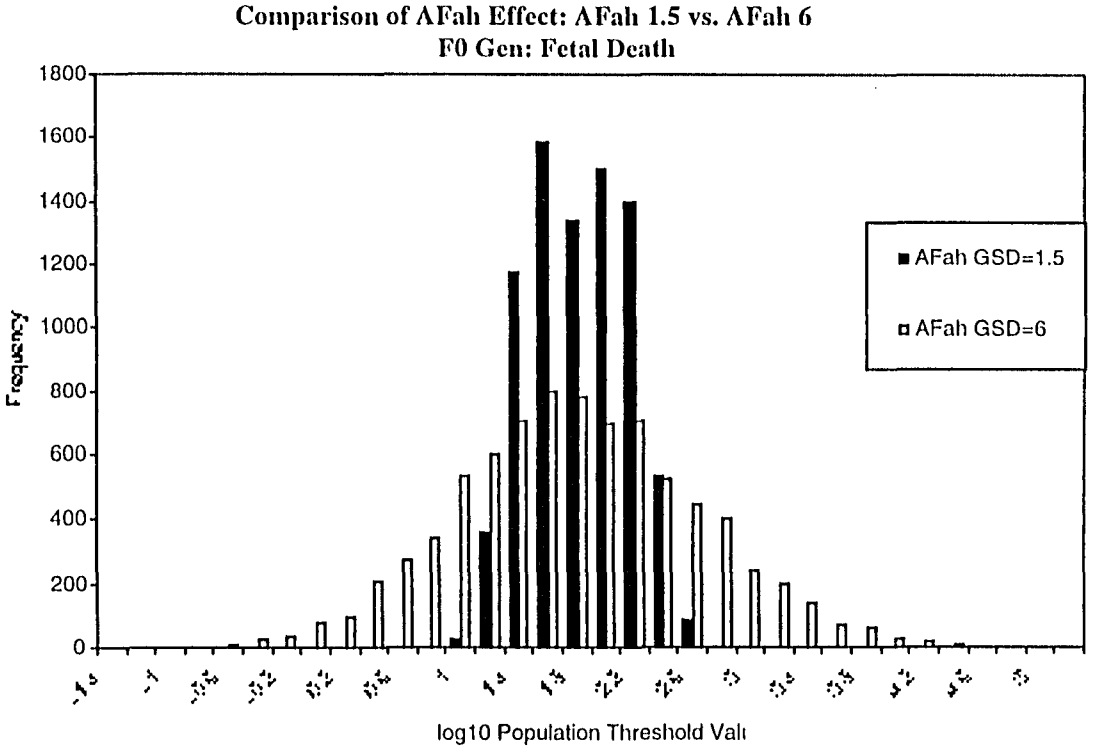
Sensitivity Analysis: Model



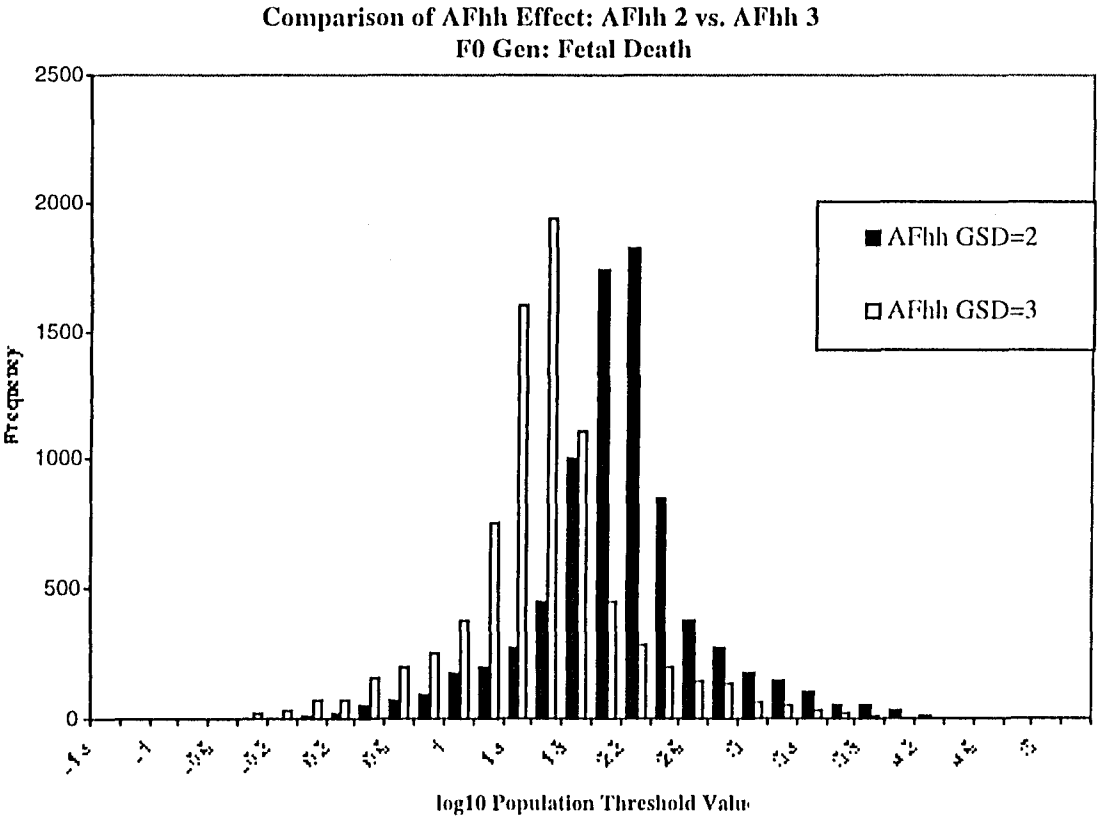
Sensitivity Analysis: Covariate



Sensitivity Analysis: AF_A



Sensitivity Analysis: AF_H



Issues not Currently Addressed Quantitatively

- Severity of effect
- Defining adverse effect
- Concordance of endpoints across species

Benefits of D-R Based Probabilistic Method

- Distribution of probability of a health impact occurring
- Risk to specified sensitive population is estimated
- Uncertainty in risk estimate is quantitatively characterized
- Level of protection is determined at end of process
- Estimate risk above and below RfD

Benefits of D-R Based Probabilistic Method

- Provides a framework for each component of extrapolation
- Components can be updated with chemical specific data
- Components contributing the greatest uncertainty can be identified and resources allocated to reduce uncertainty

Summary

- Maximizes use of available data
- Assumptions are more transparent
- Estimates risk and uncertainty in risk for specified sensitive human population
- Can be incorporated into benefit - cost analysis

References

Baird, S.J.S., Cohen, J.T., Graham, J.D., Shlyakhter, A.I., and Evans, J.S. (1996) Noncancer risk assessment: A probabilistic alternative to current practice*Human and Ecological Risk Assessment* **2(1)** , 79-102.

Baird, S.J.S., Slob, W., and Jarabek, A.M. (2000) Probabilistic noncancer risk estimation. *In preparation.*

Brand, K.P., Rhomberg, L., and Evans, J.S. (1999) Estimating noncancer uncertainty factors: are ratios of NOAELs informative?*Risk Analysis,* **19(2)** , 295-308.

Chun, J.S., and Neeper-Bradley, T.L. (1993) Two-generation reproduction study of inhaled ethylene oxide vapor in CD rats. Bushy Run Research Center, Export, PA. Study sponsored by ARC Chemical Division (Balchem), PRAXAIR, Inc. and Union Carbide Chemicals and Plastics Company, Inc.

Evans, J.S., and Baird, S.J.S. (1998) Accounting for missing data in noncancer risk assessment. *Human and Ecological Risk Assessment* **4(2)** , 291-317.

Evans, J.S., Rhomberg, L.R., Williams, P.L., Wilson, A.W., and Baird, S.J.S. (2000) Reproductive and developmental risks from ethylene oxide: A probabilistic characterization of possible regulatory thresholds. *Submitted.*

Hattis, D., Banati, P., and Goble, R. (1999) Distributions of individual susceptibility among humans for toxic effects. How much protection does the traditional tenfold factor provide for wh fraction of which kinds of chemicals and effects?*Annals of the New York Academy of Sciences,* **895** , 286-316.

Renwick, A.G., and Lazarus, N.R. (1998) Human variability and noncancer risk assessment – an analysis of the default uncertainty factor. *Regulatory Toxicology and Pharmacology,* **27** , 3-20.

Rhomberg, L.R., and Wolff, S.K. (1998) Empirical scaling of single oral lethal doses across mammalian species based on a large database. *Risk Analysis,* **18(6)** , 741-753.

Ryan, L. M. (1992) The use of generalized estimating equations for risk assessment in developmental toxicity. *Risk Analysis* , **12** , 439-447.

Schmidt, C.W., Gillis, C.A., Keenan, R.E., and Price, P.S. (1997) Characterizing inter-chemical variation in the interspecies uncertainty factor (U_H). *Fundamental and Applied Toxicology, Supplement – The Toxicologist,* **36(1, Part 2)** , 208.

Snellings, W.M., Maronpot, R.R., Zelenak, J.P., and Laffoon, C.P. (1982a) Teratology study in Fischer 344 rats exposed to ethylene oxide by inhalation. *Toxicology and Applied Pharmacology,* **64** , 476-481.

Swartout, J.C., Price, P.S., Dourson, M.L., Carlson-Lynch, H.L., and Keenan, R.E. (1998) A probabilistic framework for the reference dose (probabilistic RfD). *Risk Analysis* , **18(3)** , 271-282.

U.S. EPA (1992) Draft report: A cross-species scaling factor for carcinogen risk assessment based on equivalence of $mg/kg^{3/4}/day$. *Federal Register;* **57(109),** 24152-24173, June 5, 1992.

U.S. EPA (1994) Methods for Derivation of Inhalation Reference Concentration and Applicatior of Inhalation Dosimetry, EPA/600/8-90/066, Office of Research and Development, U.S. EPA, Washington, DC.

Appendix G

Presentation Overheads

Paul S. Price
Ogden Environmental and Energy Services

Characterizing Risks Above the Reference Dose

Presentation at the

Colloquium on Approaches to Quantifying Health Risks for Threshold or Nonlinear Effects

Paul S. Price, M.S.

Washington DC

September 28, 2000

OGDEN *ENVIRONMENTAL AND ENERGY SERVICES*

Westford, MA and Portland, ME

Topics

- Background
- Placing the RfD into a dose response framework
- Defining the "uncertainty" in the RfD
- Placing the inter- and intra- species uncertainty into a dose response framework
- Assessing risks above the RfD -two approaches
- Implications for assessing carcinogenic risks

OGDEN

Background

- Current system for noncancer risk consists of
  - A methodology for setting a "permitted dose"
  - Several closely related systems for comparing estimated dose to the "permitted dose" (HQ, MOE, Risk Cup, etc.)
- In contrast to the characterization of cancer risks no estimate of the risk (likelihood of response) associated with any dose or combination of doses
- No guidance on the meaning of doses in excess of the "permitted dose"
- Provides no guidance for the determination of benefits from reduction

Project Background

- 1993-1998 Cooperative Research and Development Agreement between EPA and private industry
- Goal: To characterize the uncertainty and variation in the assessment of noncancer risk
- Four publications: Swartout et al. 1998, Price et al. 1997 , Carlson Lynch et al. 1999, and Price et al. (Submitted-2000)
- Approach for quantitative non-cancer risks has been applied to PCBs and mercury

OGDEN

What is the RfD?

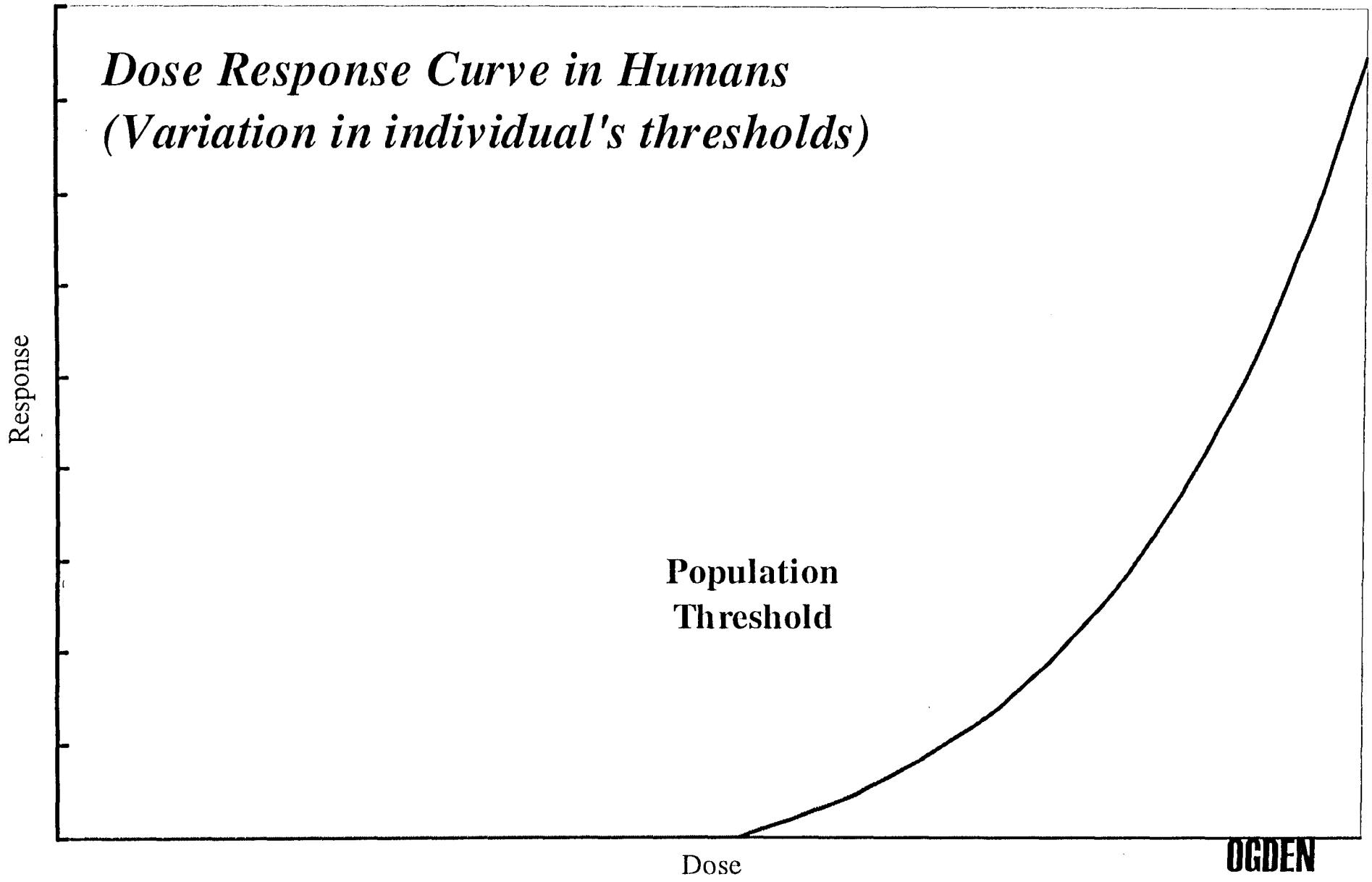
- **A policy finding:**
  - **The result of a consensus between appropriate experts after the review of a data base**
  - **A product of a political process**
- **A scientific finding:**
  - **A sub-threshold dose**
  - **A dose that is without appreciable risk**
- **The second position is assumed to be true**

OGDEN

Framework

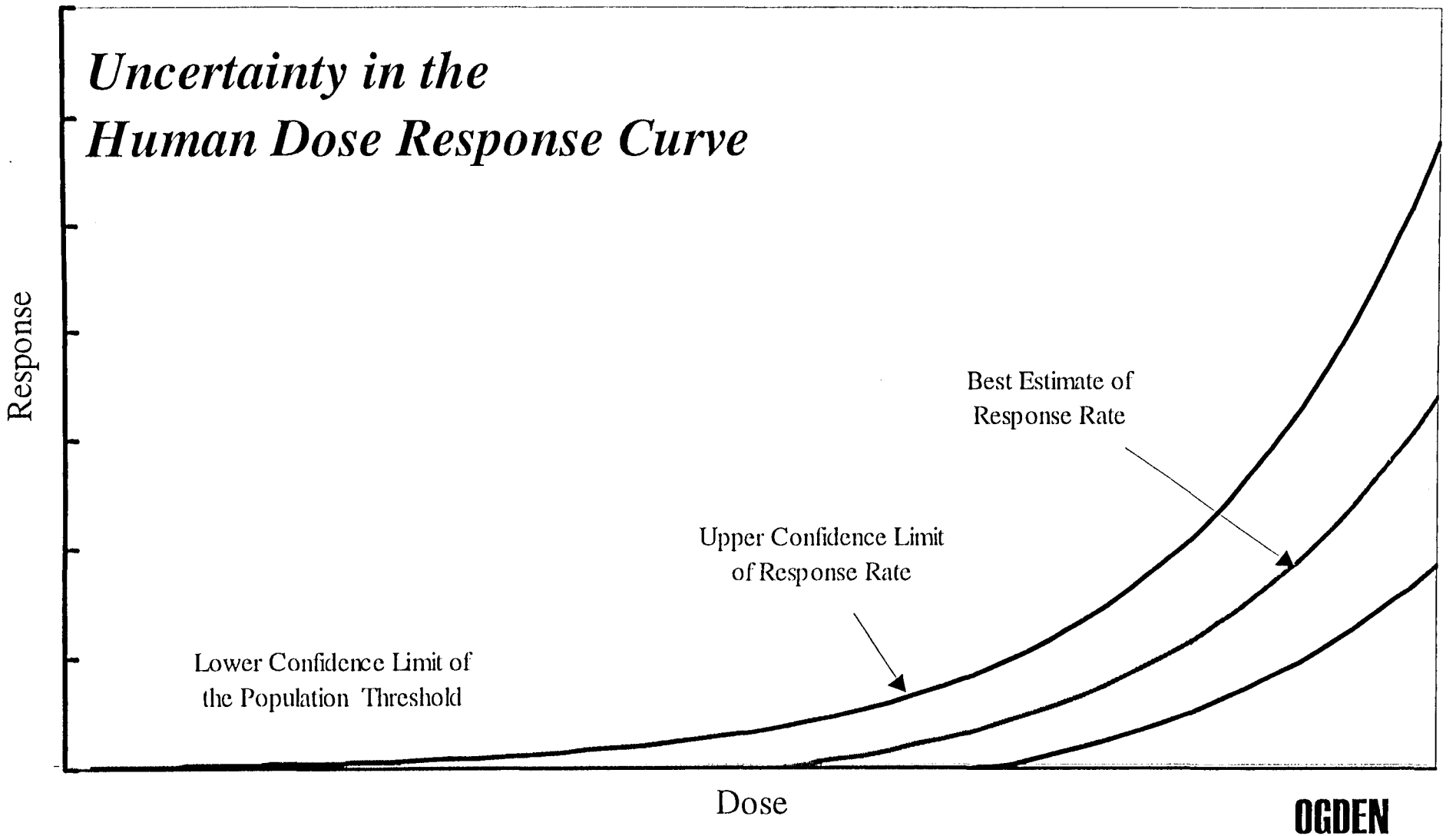
- It is important to understand how uncertainty and variation are defined in any proposed framework
- Variation refers to the level of protection
  - Fraction of the population or
  - Amount of interindividual variation
- Uncertainty comes from measurement error and interspecies extrapolation
- Toxicological criteria such as the RfD's only have uncertainty

OGDEN


Dose Response Curve in Humans
(Variation in individual's thresholds)
Response
Population
Threshold
Dose
OGDEN

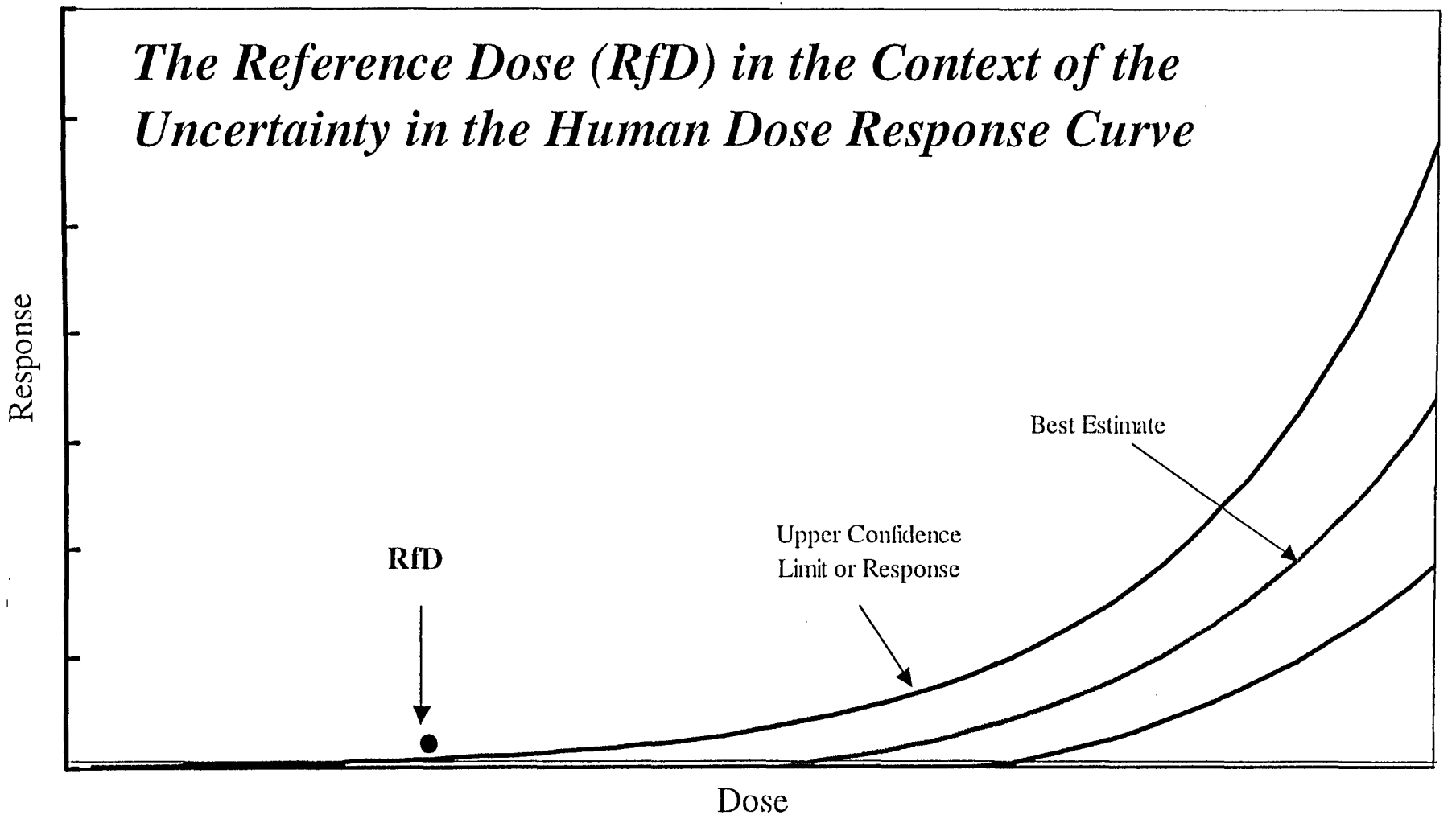
Uncertainty

- **It is rarely possible to directly determine dose-response rates for adverse effects in humans**
- **Use of mathematical models, animal surrogates, or measurements of non-adverse effects can provide a basis for estimating the curve**
- **Such estimates are subject to uncertainty**


Uncertainty in the
Human Dose Response Curve
Response
Best Estimate of
Response Rate
Upper Confidence Limit
of Response Rate
Lower Confidence Limit of
the Population Threshold
Dose
OGDEN

Proposed Definition of the RfD

- **Not an estimate of the population threshold of a compound in humans**
- **An estimate of the lower confidence limit of the threshold in humans**

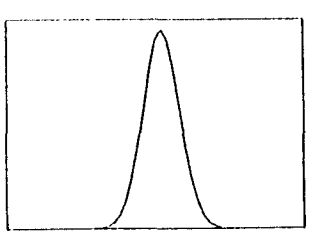

The Reference Dose (RfD) in the Context of the Uncertainty in the Human Dose Response Curve
Response
Best Estimate
Upper Confidence Limit or Response
RfD
Dose

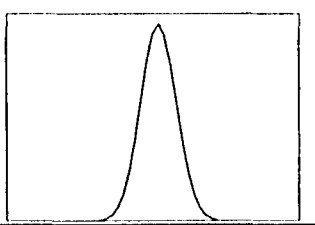
OGDEN

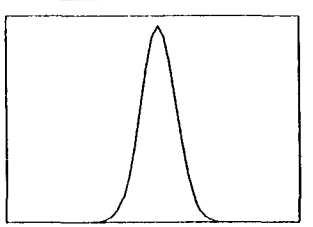
Monte Carlo Modeling and the RfD

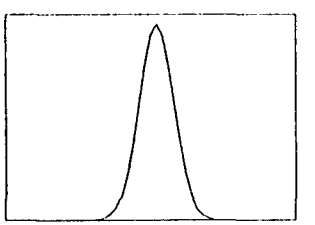
$$\text{RfD} = \frac{\text{NOAEL}}{UF_A \times UF_H}$$

$$? = \frac{\text{NOAEL}}{UF_A \times UF_H}$$

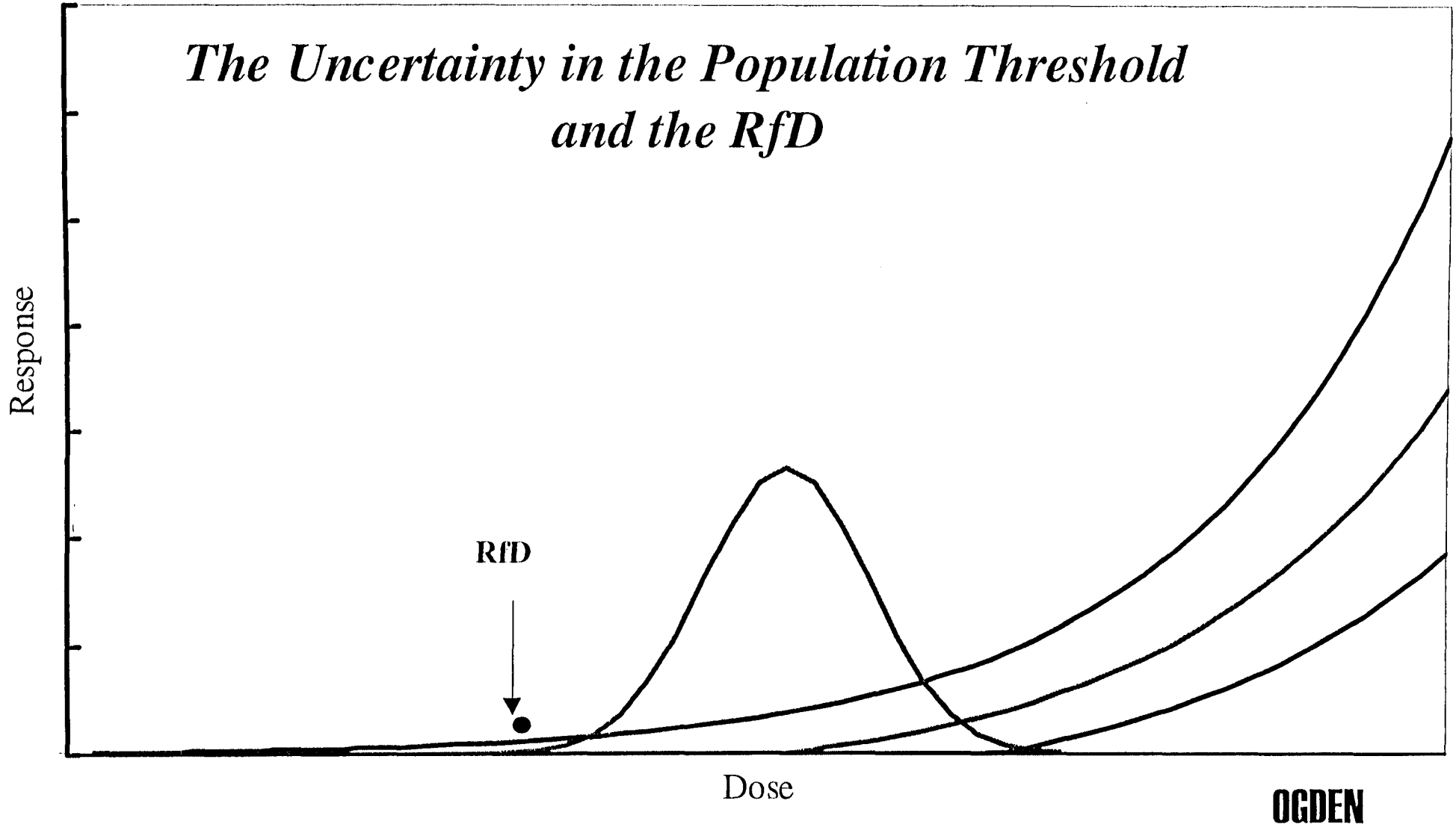




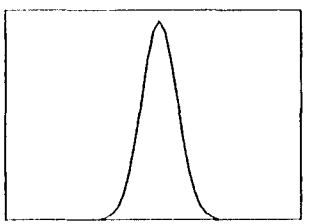


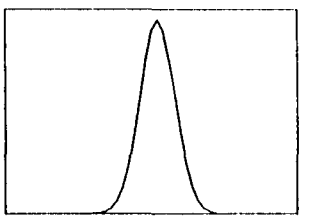


OGDEN

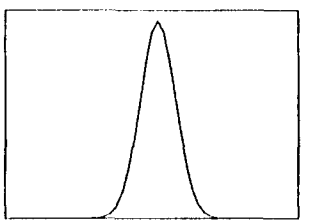
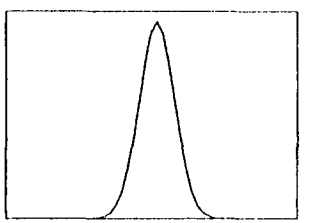

The Uncertainty in the Population Threshold and the RfD
Response
RfD
Dose
OGDEN

The Uncertainty in the Population Threshold

Population Threshold = 



NOAEL

UF_A  *x* **UF_H** 

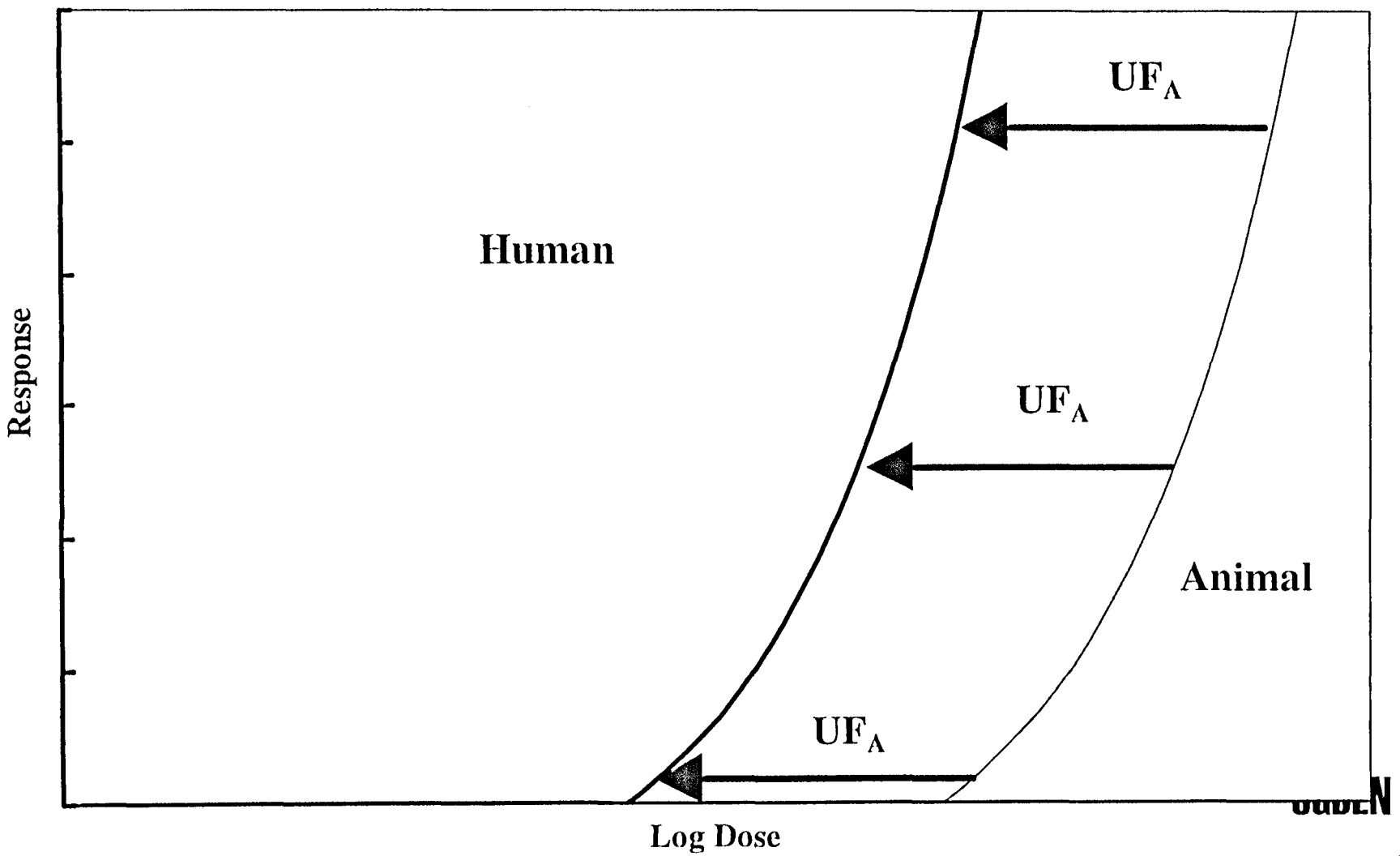
The RfD is some lower percentile of the population threshold generated by the equation.

OGDEN

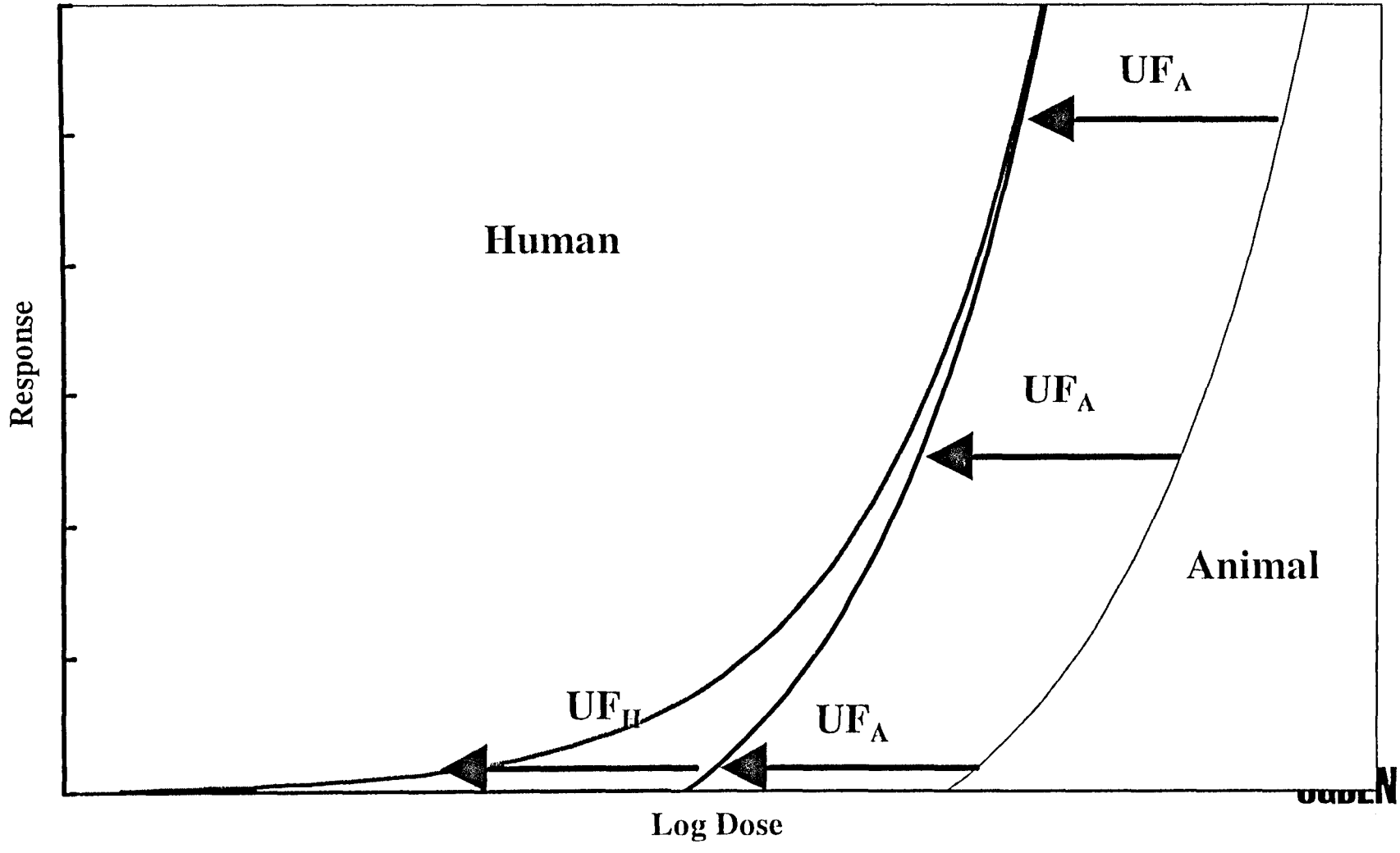
Understanding Uncertainty Factors

- Uncertainty factors can be divided into three groups
  - Primary factors (inter- and intra- species uncertainty)
  - Secondary factors (sub-chronic to chronic, LOAEL to NOAEL, database)
  - FQPA and modifying factors
- Both inter- and intra- uncertainty factors can be defined in terms of the difference between the test species and humans
  - Interspecies reflects the difference between the average sensitivity of the two species
  - Intraspecies reflects the fact that the animal model will have less inter-individual variation than humans

The Interspecies Factor and its Impact on the Extrapolation of An Animal Dose Response To Humans



The Intraspecies Factor and its Impact on the Extrapolation of An Animal Dose Response To Humans



Characterizing Risks Above the RfD

- **Two approaches can be defined**
- **The first is the general approach**
  - **Based on the proposed definitions of the RfD and the primary uncertainty factors**
  - **This approach presumes that the the shape of the dose response curve in animals is relevant to the prediction of the dose response curve in humans**
- **The second approach (minimalistic) assumes that only the estimate of the ED_{50} and NOAEL are relevant**

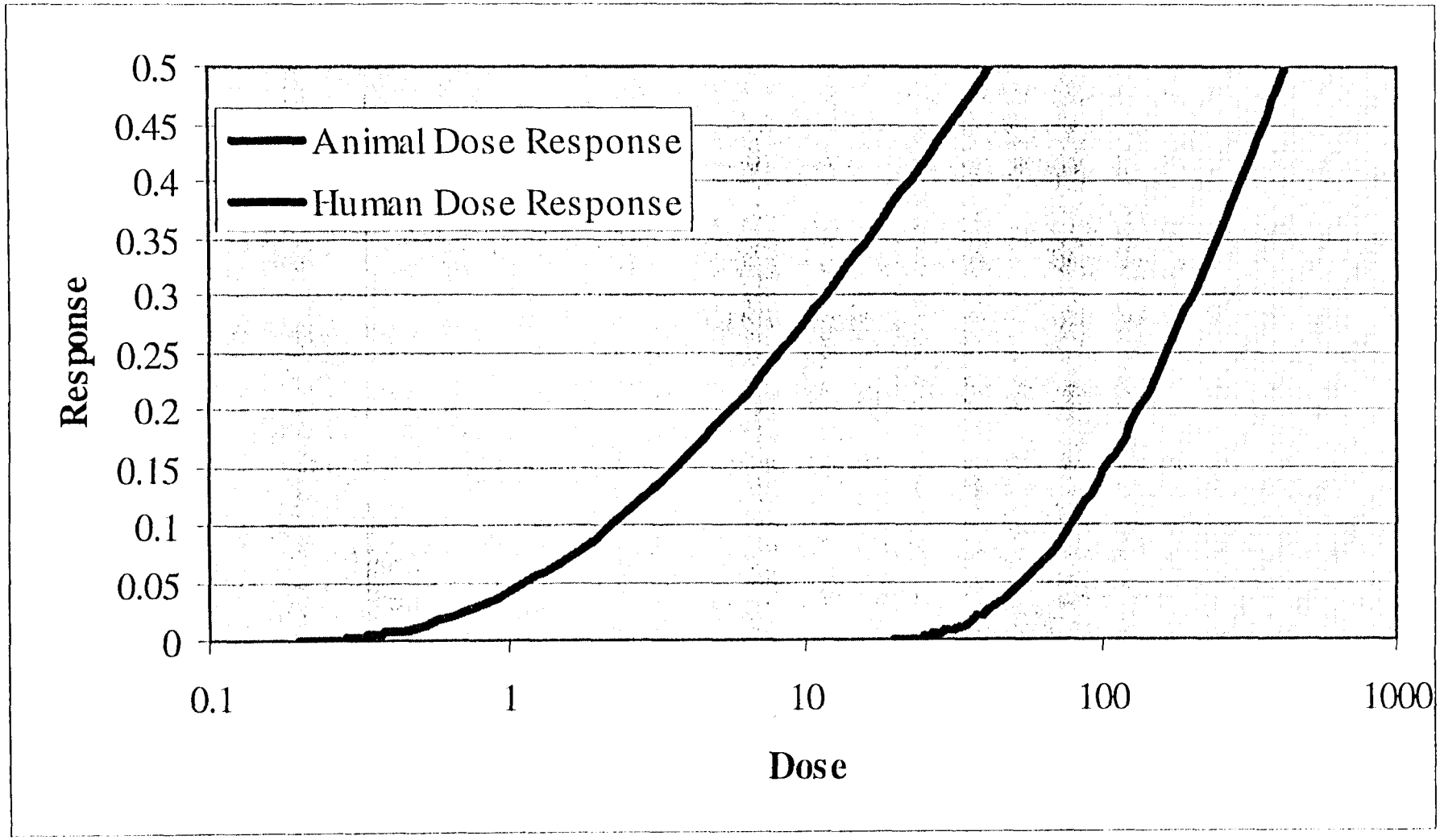
Characterizing Risks Above the RfD

- **Using the above definition of the inter and intra individual uncertainty factors it is possible to derive an equation that maps the observed dose response in animals to humans**
- **By simple algebra the relationship between the dose causing a response rate "r" in humans ED_{Ra} and the dose causing a response rate "r" in humans ED_{Rh} is given by the equation:**

$$ED_{Rh} = (ED_{50a}/UF_A)\,(ED_{Rd}/ED_{50a})^{(1-(\log UFH\ /\log(ED0a/ED50a))}$$

OGDEN

Example Extrapolation of An Animal Dose Response To Humans



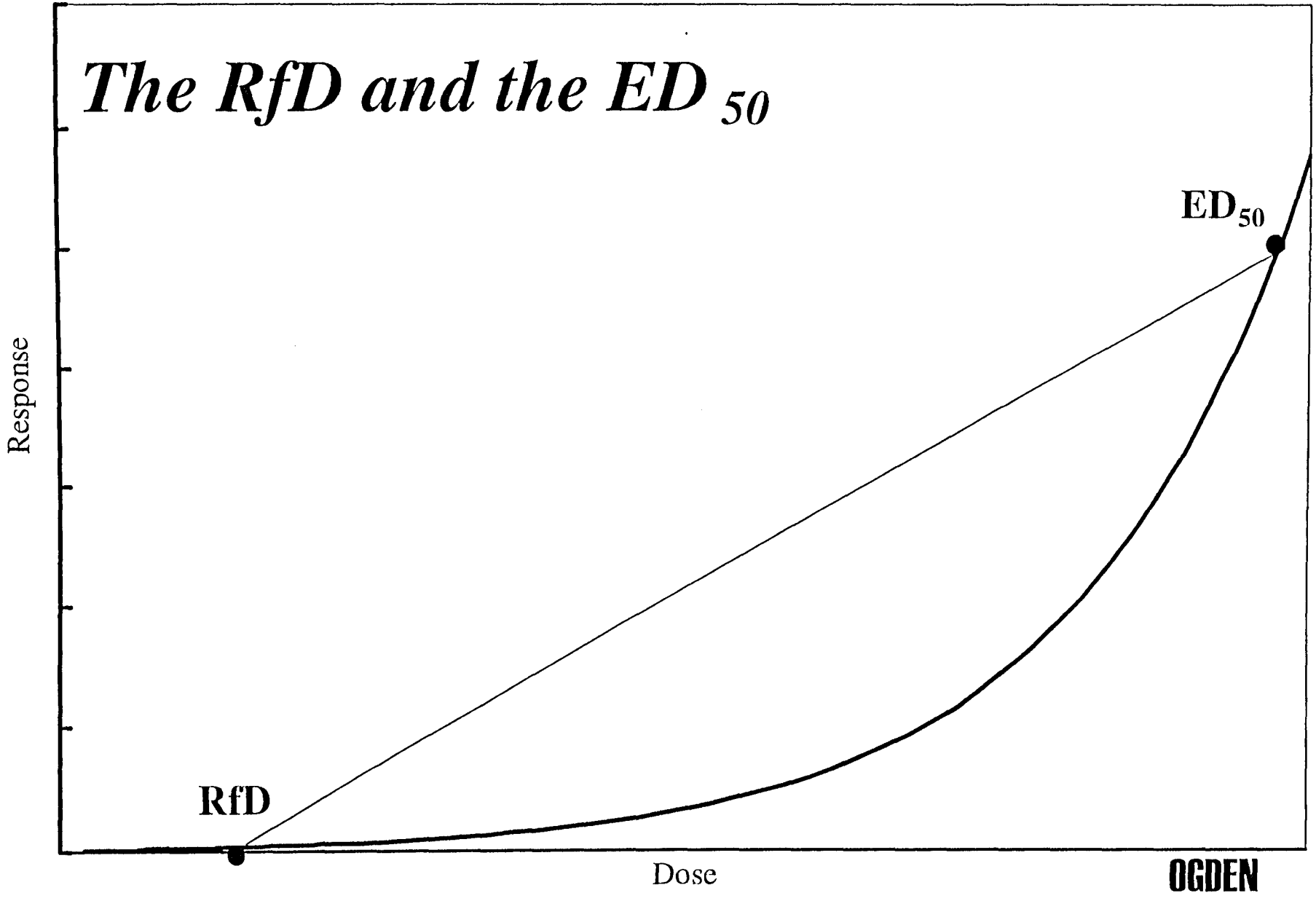
Characterizing Risks Above the RfD General Model

- **The approach assumes that the dose response curve in animals is relevant to humans**
- **Can be used in simple Monte Carlo models of uncertainty in the value of ED_{Rh}**
  - **Uncertainty in the size of the uncertainty factor required for the compound**
  - **Uncertainty in the value of ED_{Rh} (benchmark data)**
- **Requires an estimate of the threshold (ED_{0a} in the test animals**
- **Requires explicit modeling of the correlation between the values of ED_{0a} and ED_{Rh} (A topic for future research)**

Characterizing Risks Above the RfD (Minimal Model)

- **Three concepts:**
  - **View the RfD as a conservative estimate of the population threshold**
  - **Use the interspecies uncertainty factor to estimate the ED_{50h} in humans from the ED_{50a} in animals**
  - **Assume a linear response between the RfD and the ED_{50}**
- **The result is a hockey stick model of dose response**
- **Has the advantage of not requiring the assumption that the specific shape of the animal's dose response curve is a model of humans**

The RfD and the ED $_{50}$



Dose Response Equation

- The equation for the "hockey stick equation requires estimates of the doses associated with the:
  - Population threshold
  - ED_{50}
- The RfD can be used to estimate the population threshold or the uncertainty distribution can be used
- The ED_{50} for humans can be estimated based on the ED_{50} in the test animals

OGDEN

Dose Response Equation

$$Risk(d) = 0.5(d*UF_H*UF_A - NOAEL_a)/(ED_{50a}*UF_H - NOAEL_a)$$

Since a response cannot be greater than 1.0 or less than 0, the dose response relationship must be truncated such that:

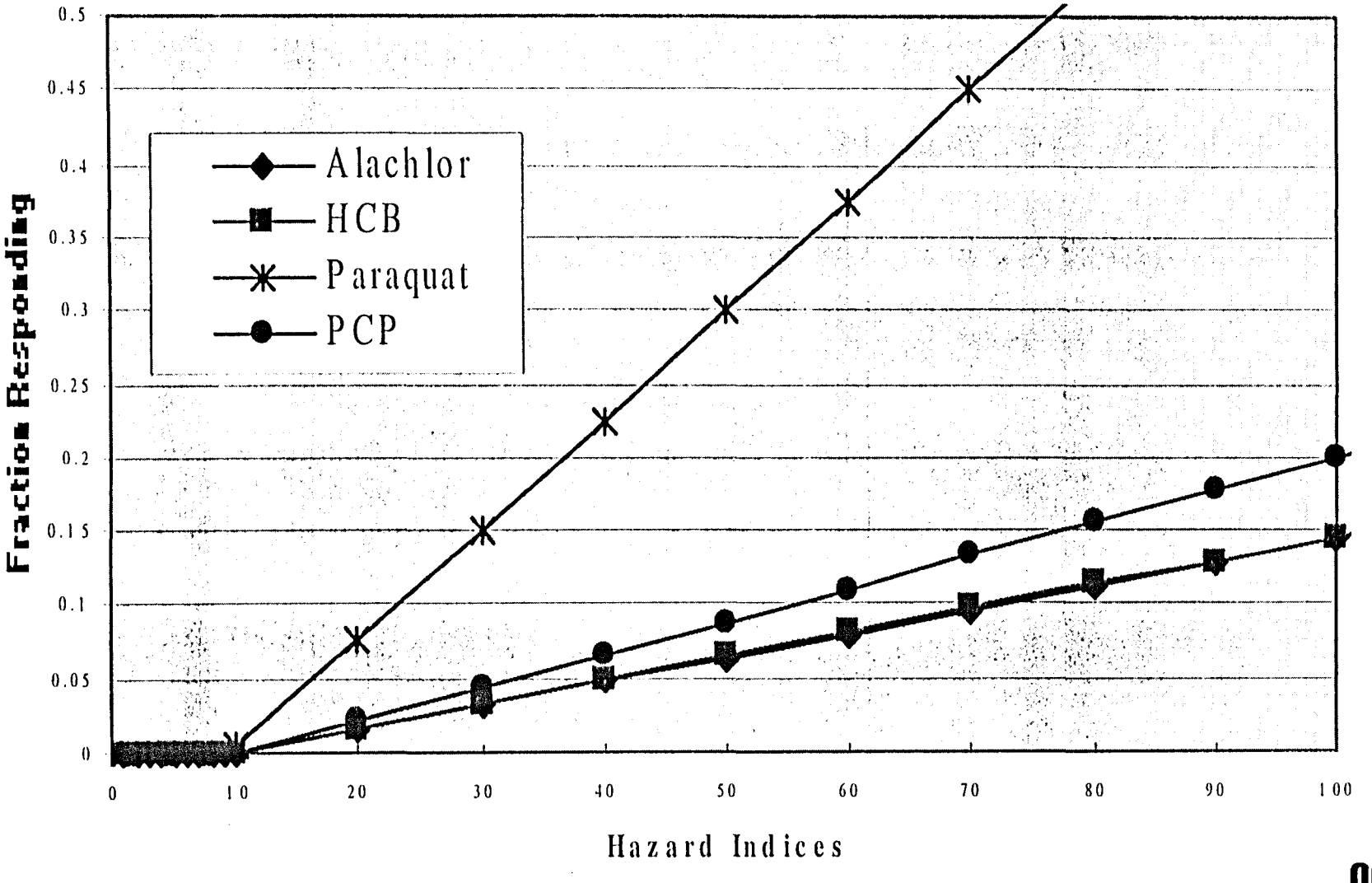
$$Risk(d) = 0 \quad \text{if } d < NOAEL_a/(UF_H*UF_A)$$

$$Risk(d) = 0.5(d*UF_H*UF_A - (NOAEL_a)/(ED_{50a}*UF_H - NOAEL_a)$$

$$Risk(d) = 1 \quad \text{if } d > 2ED_{50a} - (NOAEL_a/(UF_H*UF_A))$$

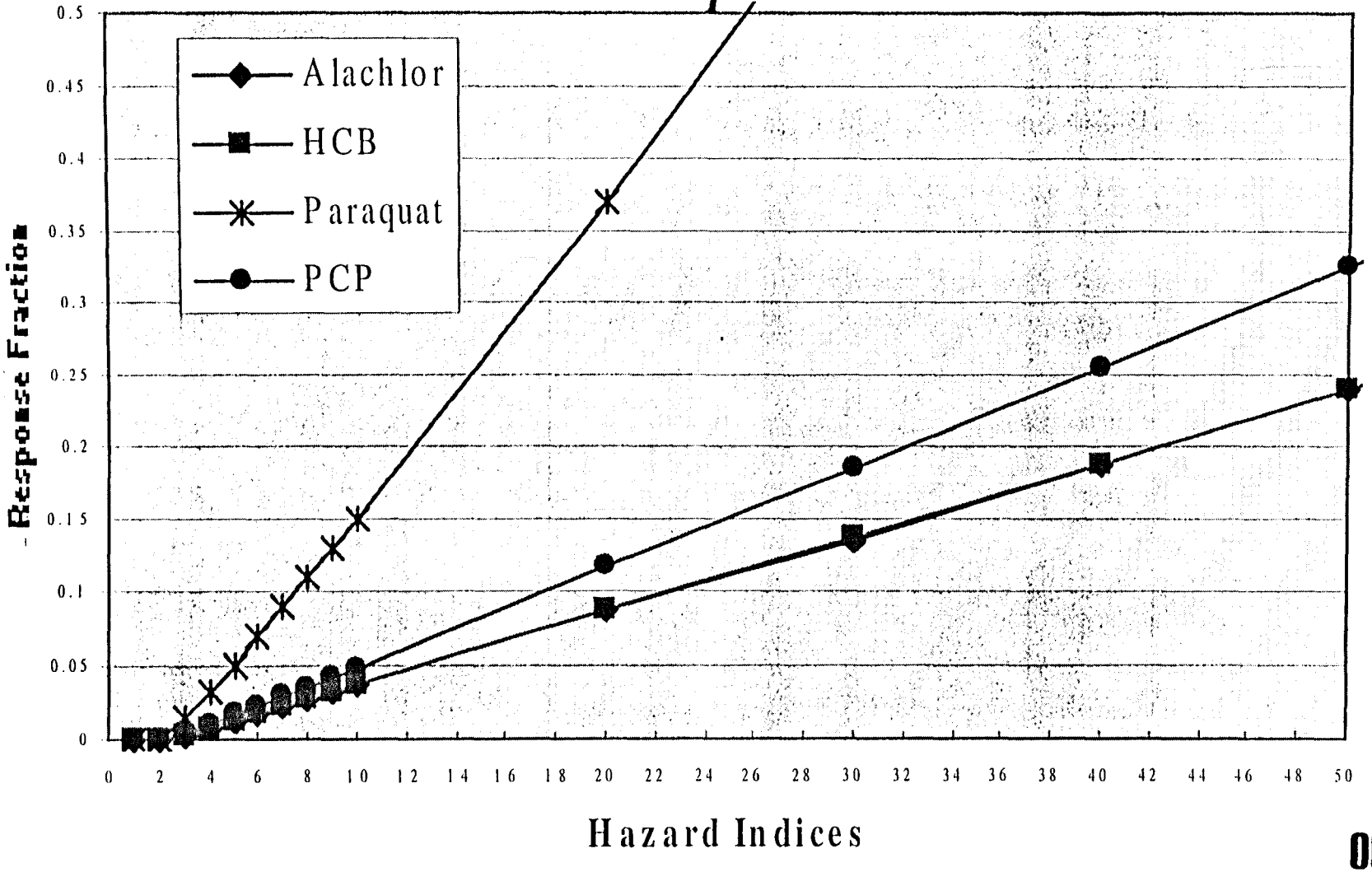
OGDEN

Unbiased Estimate of Response Rate



OGDEN

Conservative Estimates of Dose Response



OGDEN

Implications for Cancer Risk Determinations

- Either approach can be used to derive estimates of response below the dose associated with an observable response in animals
- Can use the interspecies assumptions from the cancer tradition
  - Replace the factor of 10 with a body weight based value
- Can use data on interindividual variability
  - Replace the intra individual factor with more complex functions
- Can incorporate quantitative data on uncertainty
- Approach requires some method for deriving an estimate of the threshold of carcinogenic effects

Summary

- The approach allows a quantitative estimate of risk for doses above the threshold
- Based on the establishment of a framework for the RfD and/or other criteria
- The framework is not the only possible interpretation of regulatory standards
- The approach is not limited to non-cancer
- Allows the separate modeling of variation and uncertainty
- Requires a quantitative estimate of the threshold

OGDEN

Appendix H

Presentation Overheads

Dale Hattis
Center for Technology, Environment and Development
Clark University

Expected Values of Population Dose Response Relationships

Background—Routes to Quantitative Assessment Depending on Fundamental Causal Processes and Available Types of Information

Effects Caused by Individual Threshold Processes—Homeostatic System Overwhelming

Population Dose Response Is Determined by the Population Distribution of Individual Thresholds

Data Base of Human Interindividual Variability in Pharmacokinetic and Pharmacodynamic Parameters

Observations Primarily from Pharmaceuticals

Parameters Measured Cover Various Portions of the Pathway from External Exposure to Effect

Analysis to Derive "Expected Value" Risks is Based on Observations and Assumptions About Representativeness of the Current Database, and About Distributional Forms

For individual sensitivities (among people)

For overall degrees of variability (among chemicals, controlling for types of effects and route of exposure)

Routes to Causation and Quantification of Health Effects

Direct Epidemiological Observations of Excess Health Outcomes of Concern in Relation to Exposure, After Control for Confounders

Projections Based on Changes in Intermediate Parameters Related to End Effects of Ultimate Concern (currently underdeveloped assessment methodology)

Excess Infant Mortality in Relation to Birth Weight Changes

Decreased Male Fertility in Relation to Sperm Count/Quality Changes

Increased Cardiovascular Mortality in Relation to Changes in Cardiovascular Risk Factors (FEV1, Blood Pressure, Heart Rate Variability, Serum Fibrinogen)

Projections Based on Population Distributions of Individual Susceptibility to Effects Caused by Overwhelming Homeostatic Systems (this talk)

Projections Based on Incremental Addition to Stochastic Background Mutation Processes (e.g. Primary Genetic Mechanisms of Carcinogenesis)

Data Base of Human Interindividual Variability in Pharmacokinetic and Pharmacodynamic Parameters

Observations Primarily from Pharmaceuticals

Data Sets Selected Provide Individual Data for at Least 5 Reasonably Healthy People

Parameters Measured Cover Various Portions of the Pathway from External Exposure to Effect

Current Data Base has 443 Total Data Groups (Each Yielding a Variability Observation)

- 11 Contact Rate (2 for children)
- 343 Pharmacokinetic (71 include children)
- 89 with Pharmacodynamic (and often also pharmacokinetic) Information (6 include children)

Variability is Predominantly Lognormal—Expressed as Log(GSD)—the standard deviation of the Logs of the primary data points

Within Specific Data Types, Distributions of the Log(GSD)'s Themselves Are Reasonably Close to Lognormal.

Challenges for Modeling Human Variability in Susceptibility

Diverse Data Types--each provides information about interindividual variability for a portion of the pathway from exposure to effect

--Characterize each data type with "dummy" (0,1) variables to represent the presence or absence of variability due to each step in the causal pathway

Form of the Distribution(s) of Human Interindividual Variability for Different Parameters

--Assume lognormality

--Combine variability from multiple causal steps by adding together lognormal variances

Differences Among Chemicals in Amounts of Variability--Distinguishing the Spread of Variability Estimates Due to Measurement Errors from the Real Spread of Variability Among Chemicals

--Assess the spread of model predictions from observed variability for statistically stronger vs weaker observations

Components That Can Contribute to the Interindividual Variability in Different Measured Parameters

- Contact Rate (Breathing rates/body weight; fish consumption/body weight)
- Uptake or Absorption (mg/kg)/Intake or Contact Rate
- General Systemic Availability Net of First Pass Elimination
- Dilution via Distribution Volume
- Systemic Elimination/Clearance or Half Life
- Active Site Availability/General Systemic Availability
- Physiological Parameter Change/Active Site Availability
- Functional Reserve Capacity--Change in Baseline Physiological Parameter Needed to Pass a Criterion of Abnormal Function

Basic Methodology for Assessing Expected Values for the Incidence of Individual Threshold Responses as a Function of Dose

1. Use the human database to make central estimates of overall lognormal variability [as a log(GSD)] from the observed variances associated with various causal steps—depending on the route of exposure, the type of effect, and the severity of the response to be modeled.

2. Determine the lognormal uncertainty in log(GSD)'s estimated from the model for the largest data sets—reducing the inflating influence of statistical sampling errors on the observed spread of log(GSD) estimates for individual cases.

3. Sample repeatedly from the assessed lognormal distribution of log(GSD) values, and calculate arithmetic average of risks for people exposed at various fractions of the dose causing a 5% incidence of effect in humans (for model calculations this is done in a simple Excel spreadsheet, without the need for a formal Monte Carlo simulation model).

4. Summarize the results as simple power-law functions.

Additional Challenges--Not Addressed in the Current Analysis

Unrepresentativeness of the Populations Studied --Children, Elderly, and Sick Likely to be Underrepresented (although some appreciable children's data is included in the current data base) (leads to some understatement of likely variability)

Inclusion of Measurement Errors in the Basic Observations as Part of Apparent Variability (leads to some overstatement of likely variability)

Possible Unrepresentativeness of the Chemicals Studied --too many "problem chemicals" and "problem responses" with more variability than might be seen by agencies dealing with usual drugs, food additives?

A Scale For Understanding Lognormal Variability--Fold Differences Between Particular Percentiles of Lognormal Distributions

Log_{10} (GSD)	Probit slope [$1/Log_{10}$ (GSD)]	Geometric standard deviation	5%-95% Range (3.3 standard deviations)	1%-99% Range (4.6 standard deviations)
0.1	10	1.26	2.1 fold	2.9 fold
0.2	5	1.58	4.5 fold	8.5 fold
0.3	3.33	2.0	10 fold	25 fold
0.4	2.5	2.5	21 fold	73 fold
0.5	2	3.2	44 fold	210 fold
0.6	1.67	4.0	94 fold	620 fold
0.7	1.43	5.0	200 fold	1800 fold
0.8	1.25	6.3	430 fold	5,300 fold
0.9	1.11	7.9	910 fold	15,000 fold
1	1.0	10.0	1,900 fold	45,000 fold
1.1	0.91	12.6	4,200 fold	130,000 fold
1.2	0.83	15.8	8,900 fold	380,000 fold

Summary of Unweighted Log(GSD) Variability Observations for Different Types of Uptake and Pharmacokinetic Parameters

Parameter Type	Oral	IV	Inhaled	Other Routes	All Routes + Route-Nonspecific
Blood concentration for toxicant	.322[a] (3) .295-.351				.322 (3) .295-.351
Body weight (adults only)					.086 (2) .065-.113
Contact rate/body weight	.299 (2) .227-.393		.090 (3) ..059-.137	.168 (1)	.149 (6) ..066-.336
Volume of Distribution/body weight					.124 (49) .058-.284
Volume of Distribution with no control for body weight					.109 (5) .070 .170
Cmax/(dose/body weight)	.156 (28) .067-.362	.121 (3) .062-.237	.071 (1)	.176 (2) .113-.273	.150 (34) .067-.337
Cmax/dose with no control for body weight	.160 (12) .074-.374	.150 (2) .110-.204	.252 (1)	.227 (4) .167-.307	.175 (19) .090-.339
Elimination Half-Life or Clearance/Body Weight					.129 (136) .068-.248
Clearance with no control for body weight					.137 (5) .076-.248
AUC/(dose/body weight)	.169 (35) .084-.341	.125 (14) .075-.209	.149 (1)	.139 (5) .061-.317	.154 (55) .078-.301
AUC/dose with no control for body weight	.200 (24) .102-.391	.140 (5) .080-.246	.354 (2) .169-.742	.257 (4) .202-.327	.202 (35) .104-.391
Total uptake and pharmacokinetic observations	(106)	(24)	(6)	(16)	(354)

Ranges are approximate 10th and 90th percentiles of the individual data sets in each category.

Summary of Unweighted Log(GSD) Variability Observations for Different Types of Pharmacodynamic Parameters

	GI Tract	Nervous System	Resp. System	Cardiovascular Renal System + Receptor-Based Effects	Other (e.g., eye, skin irritation)	All Effects
Local (Contact Site) Parameter Change/External Exposure or Dose			Acute .655 (17) .369-1.16 Chronic .279 (1)			Acute .655 (17) .369-1.16 Chronic .279 (1)
Local (Contact Site) Response/External Exposure or Dose	.325 (1--stomach pH)		.475 (7) .208-1.087		.433 (8) ..227-.825	.443 (16) .221-.887
Physiological Parameter Change/Internal Concentration After Systemic Delivery		.259 (6) .200-.337		.175 (13) .072-.425	.536 (4) .330-.869 (Immune)	.235 (23) .098-.566
Physiological Parameter Change/External Systemic Dose		.235 (1)		.276 (1)		.232 (2) .170-.317
Response/Blood Level or Internal Concentration After Systemic Delivery		.247 (11) .109-.561		.297 (5) .108-.815	.060 (Immune) .502 (cataracts)	.250 (18) .097-.644
Response/External Dose (IV or Oral Admin.) Without Large Dosimetric Uncertainty		Oral .527 (2) IV .359 (3) Inhl .051 (2)		.266 (1)		.233 (8) .065-.836
Response/External Dose With Large Dosimetric Uncertainty (e.g. workplace epidemiology)			1.33 (1--talc ung disease)	.684 (3) .430-1.09		.807 (4) .456-1.43
Total Observations Including Pharmacodynamic Variability	(1)	(26)	(25)	(23)	(14)	(89)

STATISTICAL UNCERTAINTY IN THE ESTIMATES OF INTERINDIVIDUAL VARIABILITY

Weight of Each Variability Observation = 1/Variance of Log[log(GSD)]

Continuous Parameters--Empirical Formula Derived from Standard Statistical/Sampling Error Variance of Normally-Distributed Data

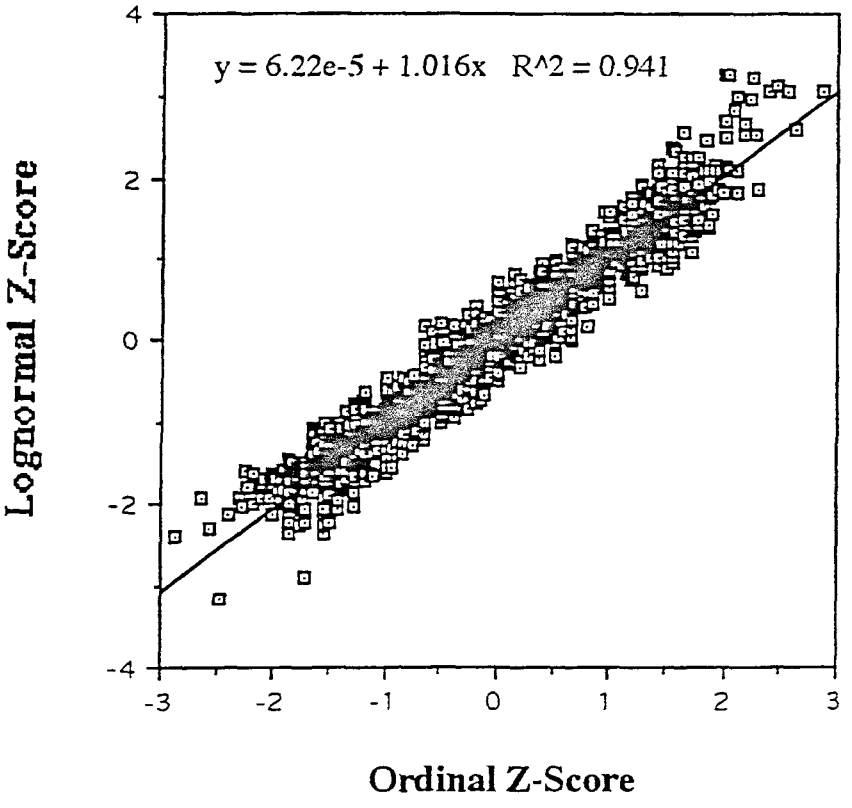
$$\text{Weight} = 10.6\ N - 10.33,$$

where N = the number of people studied

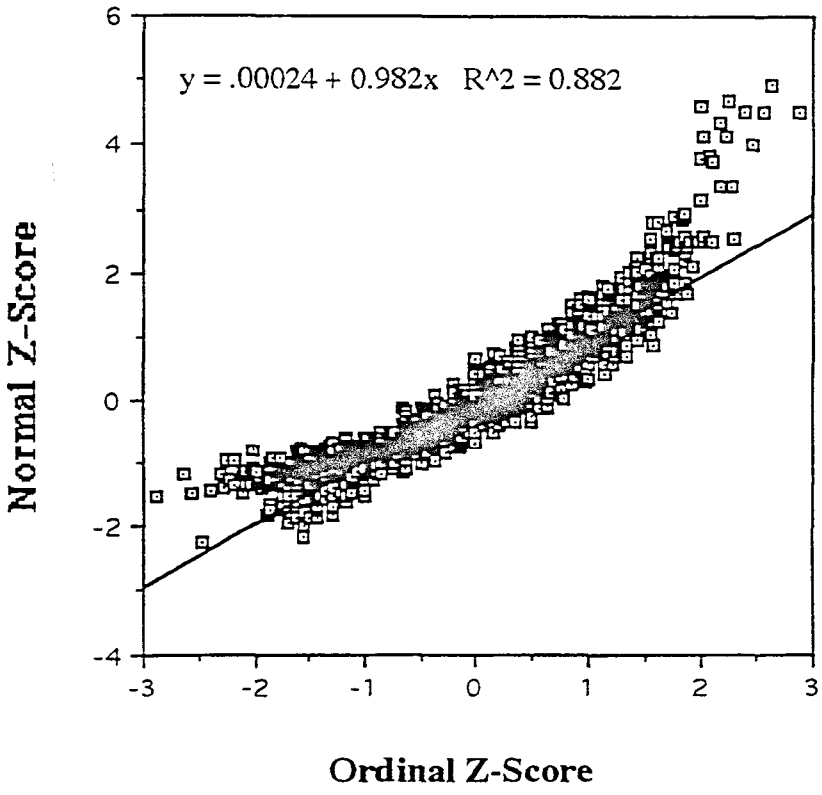
Quantal Response Parameters--Variance Derived from 10 Points of Likelihood Distribution (5th - 95th % confidences levels) Fit Using Haas Spreadsheet System

Comparison of 2700 Pharmacokinetic Data Points with Expectations Under Lornormal and Normal Distributions

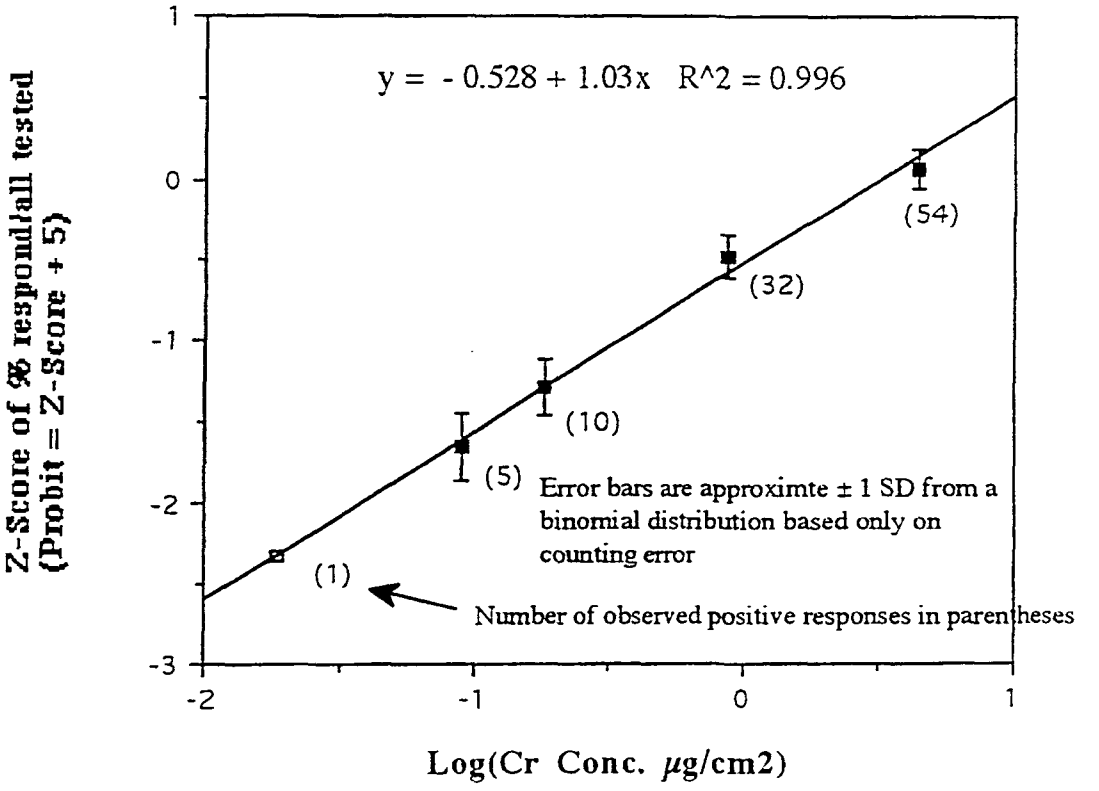
A. Lognormal Comparison



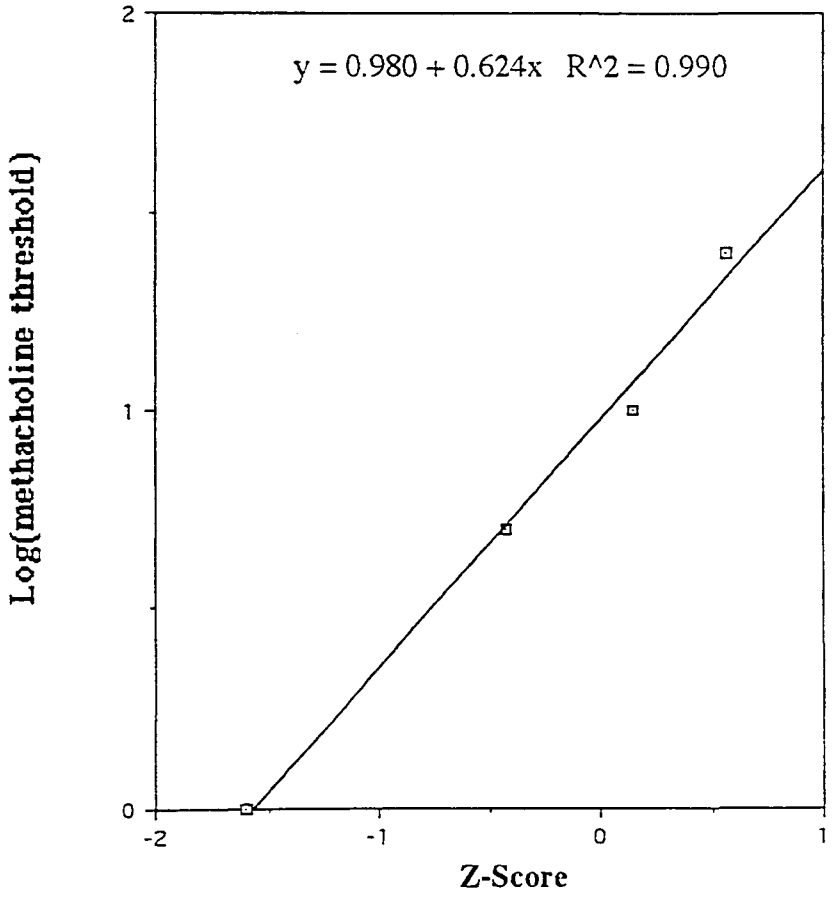
B. Normal Comparison



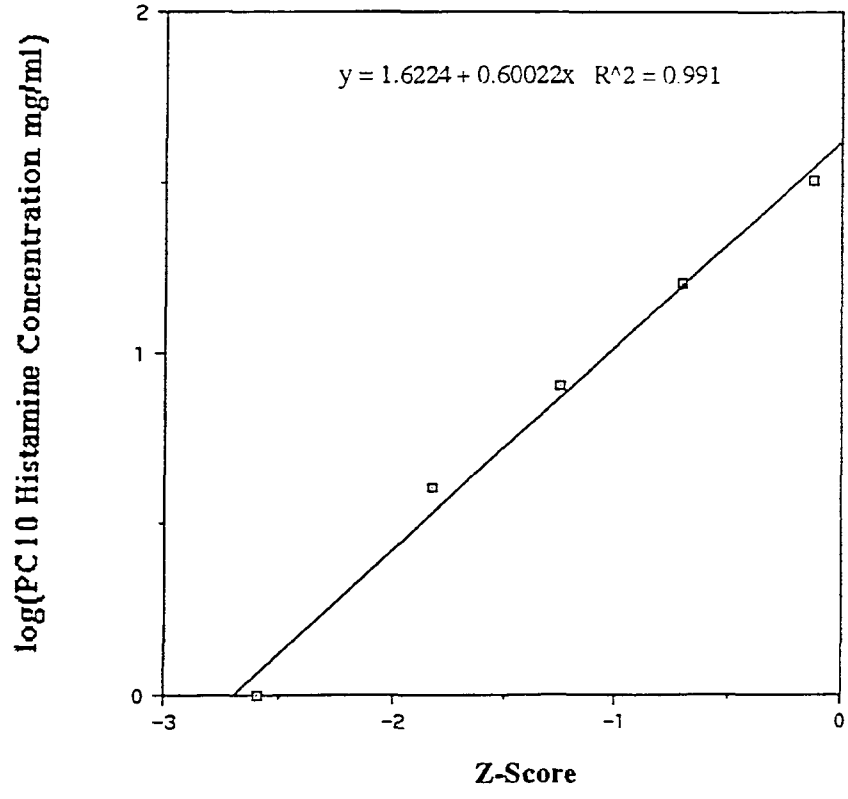
Log Probit Plot of the Percentage of 102 Tested People Who Gave Positive Skin Patch Tests for Chromium (VI)--Data of Nethercott et al. (1994)



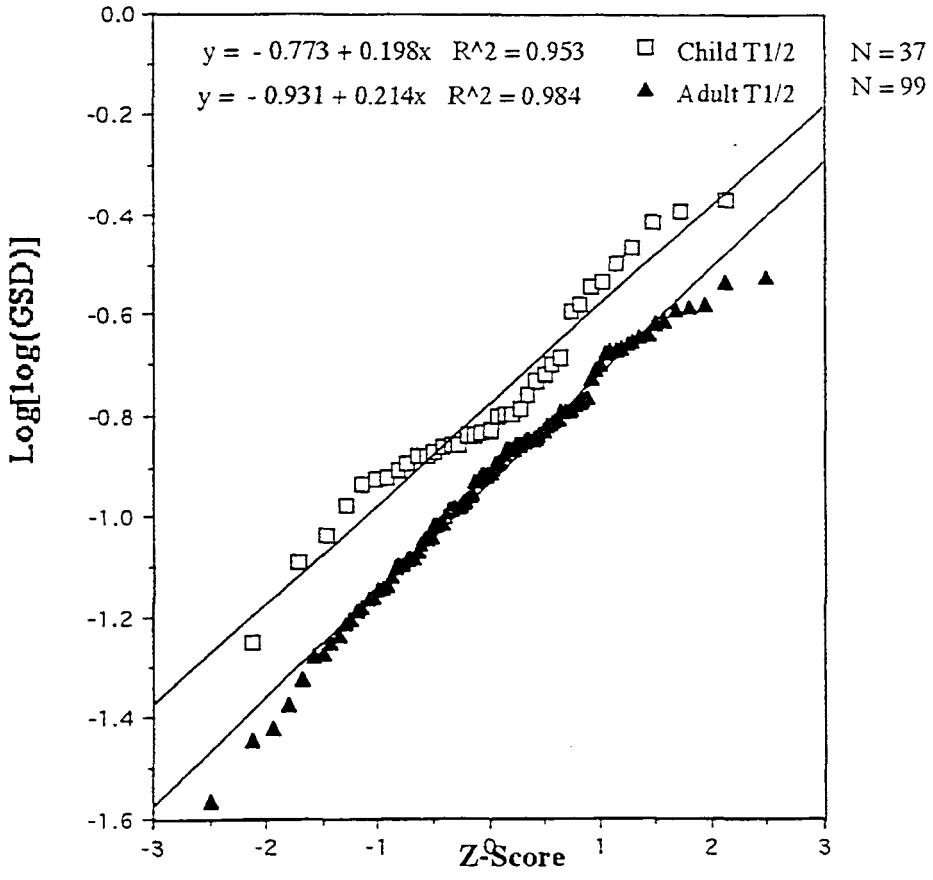
Lognormal Plot of the Distribution of PC20 Methacholine Response thresholds in 5623 Smokers with Mild to Moderate Airflow Obstruction--Data of Tashkin et al., 1996



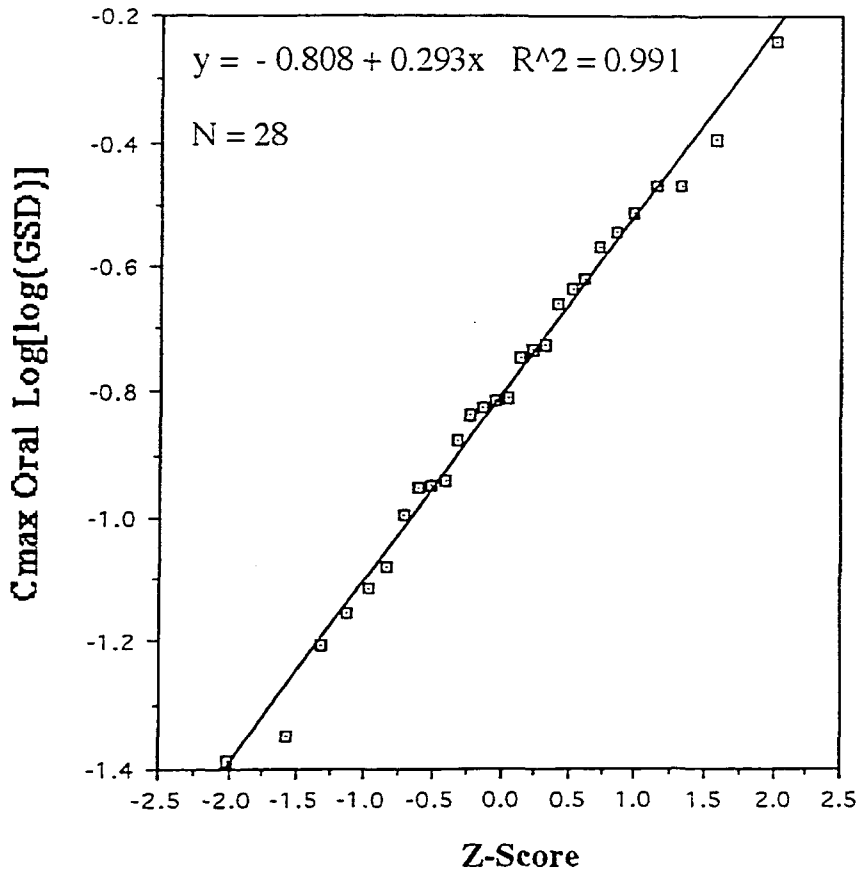
Lognormal Plot of the Distribution of PC10 Histamine Response Thresholds in 1892 Randomly Selected Adults from Two Dutch Communitites--Data of Rijcken et al. (1987)



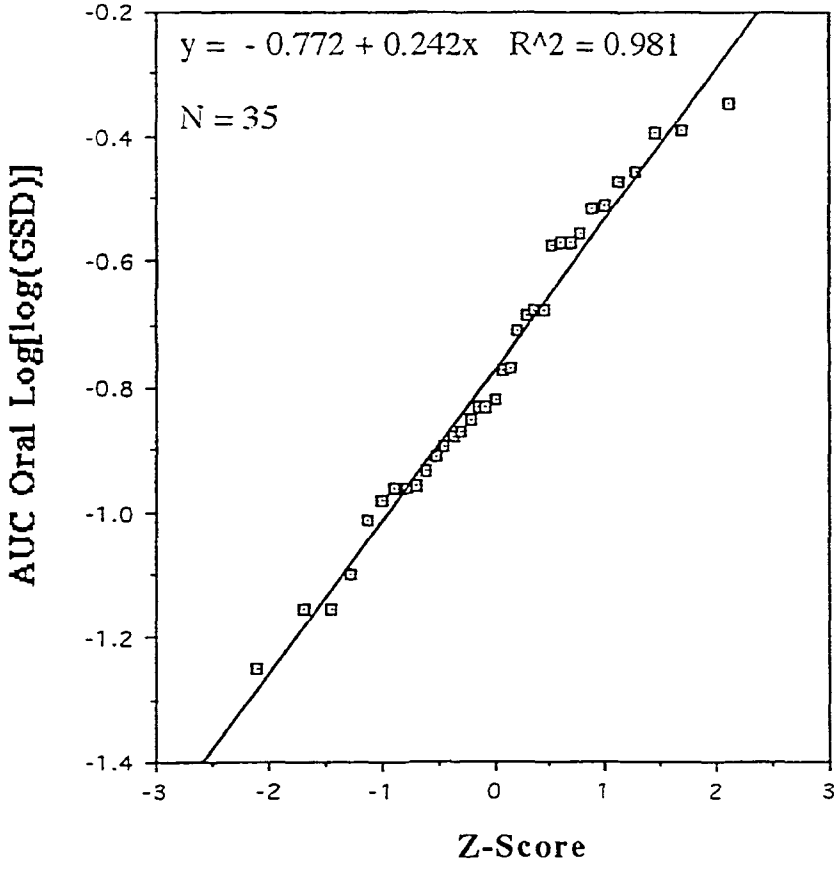
Lognormal Plots of Log(GSD) Variability Observations for Elimination Half Lives for Groups Inclucing Children (<12 years) vs Groups Including Only Adults



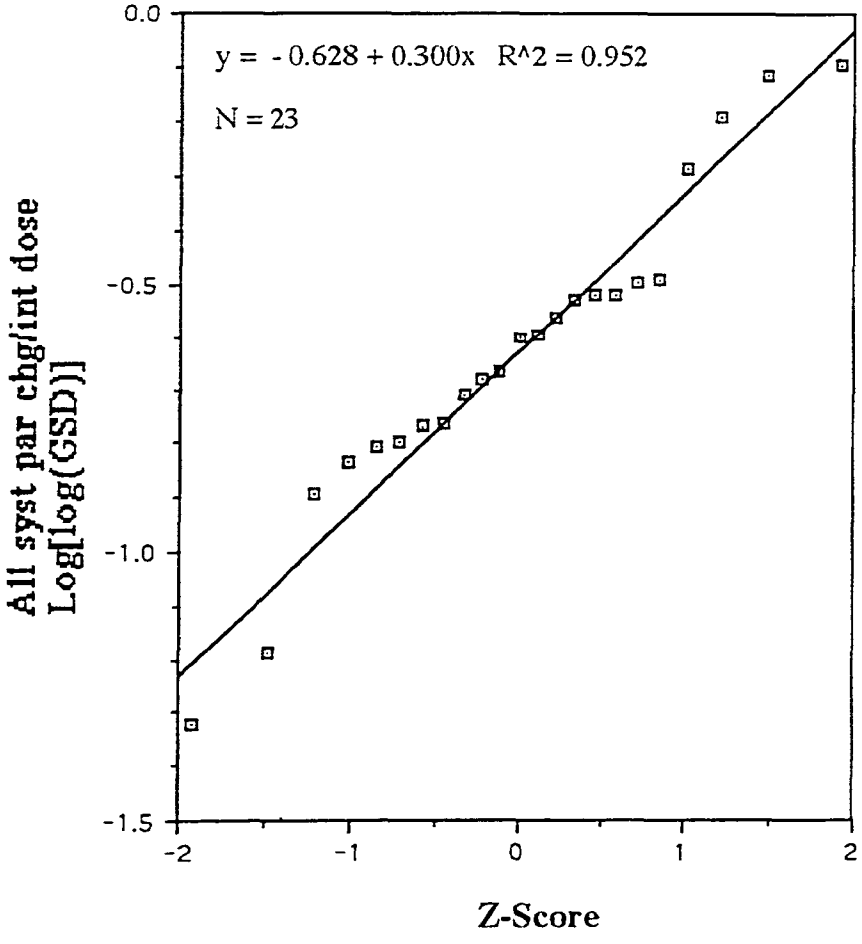
Lognormal Plot of Log(GSD) Variability Observations for Oral Cmax Values



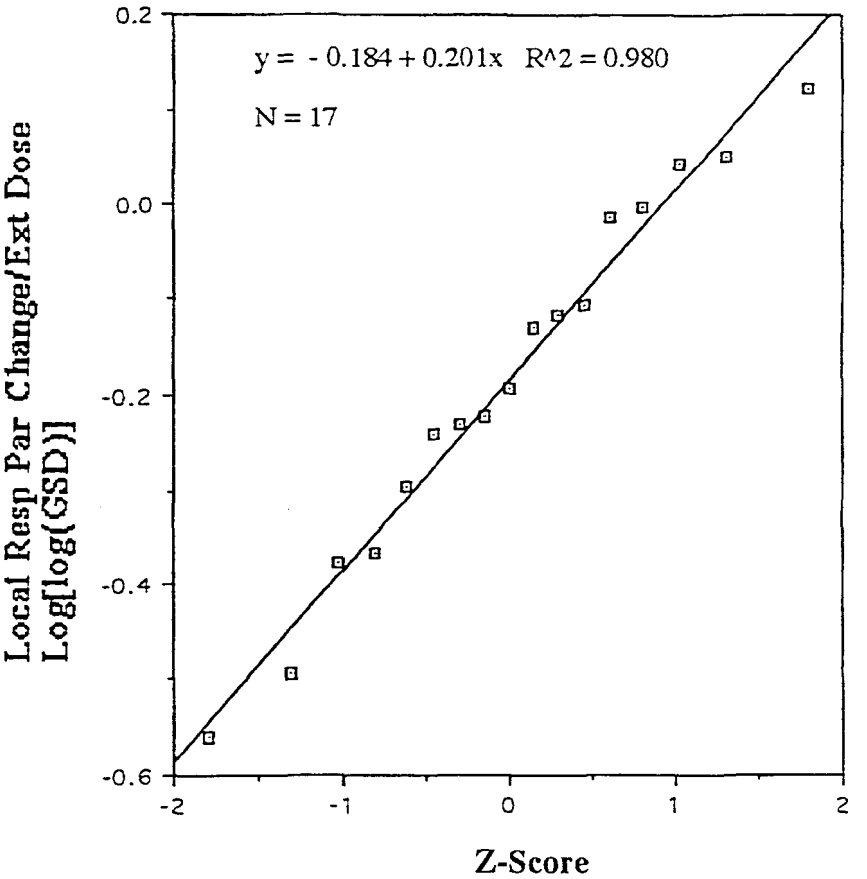
Lognormal Plot of Log(GSD) Variability Observations for Oral AUC Values



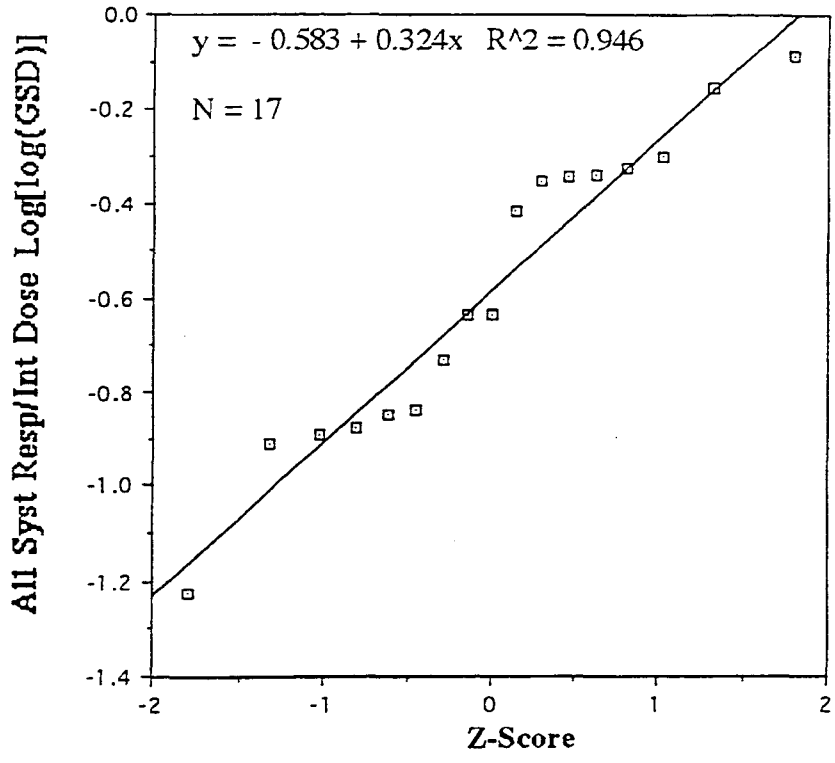
Lognormal Plot of Log(GSD) Variability Observations for All Systemic Parameter Changes In Relation to Internal Doses



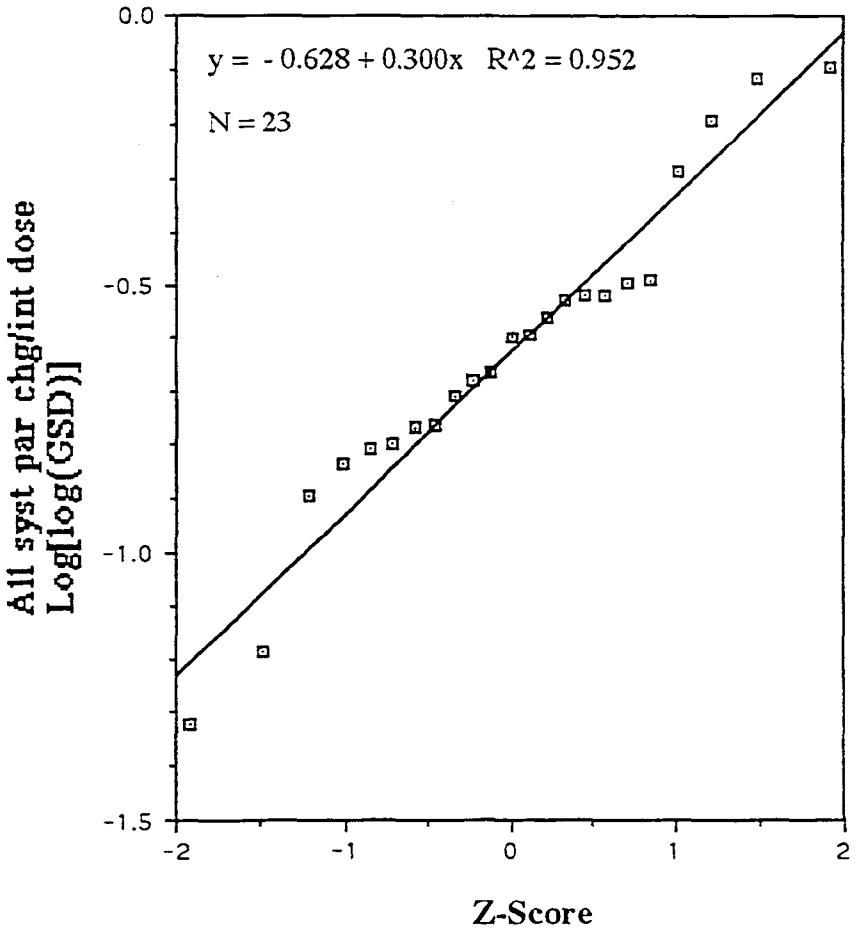
Lognormal Plog of Log(GSD) Variability Observations for Acute Changes in Lung Function in Response to External Exposure



Lognormal Plot of Log(GSD) Variability for All Systemic Responses in Relation to Internal Doses or Blood Levels



Lognormal Plot of Log(GSD) Variability Observations for All Systemic Parameter Changes In Relation to Internal Doses



Examples of Tentative "Severity" Categorizations by Type of Effect

Responses Rated as "Mild Reversible"

Olfactory cognition--air concentrations needed to produce 3 levels of smell perception

Nasal Dryness

Throat Irritation

Nose irritation--slight or moderate

Pulmonary discomfort--"slight" and "moderate" or more

Eye irritation--External air concentration causing 4 levels

Skin hypersensitivity to chromium (VI)

Skin hypersensitivity--lowest dilution of allergen needed to cause a 2mm diameter wheel

Skin irritation response to sodium laurel sulfate applied via skin patch

Eye irritation--slight or moderate and above

Paresthesia/blood level

Achievement of a specific degree of cardiac blood flow (unblocking of a clot) following an infarction in relation to the 2-90 minute AUC of a tissue plasminogen activator

Skin Rash in relation to plasma concentration

"Adequate" sedation/drowsiness

Analgesia from dental pain (not taking medication at 3 and 6 hours after procedure)

Suppression of coughing (2 levels) on intubation

Creation of conditions for intubation (2 levels--"excellent" and "good")

Responses Rated as "Moderate-Severe Reversible" or Irreversible

"Significant" hearing loss/one dose of cisplatin

Haloperidol toxicity (minimum of 4 other signs plus, in some cases seizures, catatonia, mental confusion) in relation to maximum blood level

Disarthria/blood level

Hearing defects/blood level

Visual effects/blood level

Anxiety/blood cholinesterase

Psychomotor depression/blood cholinesterase

Unusual dreams/blood cholinesterase

High ß2M urinary excretion vs occupational blood conc X time

Digoxin toxicity in relation to serum digoxin concentration

Cataracts in relation to TNT hemoglobin adducts

Dose-limiting toxicity including malaise, neurotoxicity, pericardial effusion and coagulopathy

End tidal concentration for anesthesia (not moving in response to stimulus)

Neutropenia (2 levels)

Pneumoconiosis (2 levels) in relation to cumulative talc air exposure

Responses Rated as "Severe" and/or Irreversible

Ataxia/blood level

Deaths/blood level

Deaths/red blood cell cholinesterase inhibition "hits"

Model Estimates of Human Interindividual Variability Partitioned Among Various Causal Steps

A. Pharmacokinetic Steps

	Central Estimate Log(GSD)
Total Number of Variability Data Sets Included	443
Oral Contact Rate (e.g. tap water/kg BW)	0.262
Inhalation Contact Rate (breathing rate/kg BW)	0.091
Other Contact Rate	0.168
Oral Uptake or Absorption(mg/kg)/Intake or Contact Rate	0.000
Inhalation Fraction Absorbed	0.000
Other Route Fraction Absorbed	0.000
Oral Systemic Availability Net of Local Metabolism or First Pass Liver Elimination	0.124
Systemic Availability After Absorption by Inhalation or Other Route	0.147
Body weight correction	0.086
Dilution via Distribution Volume/BW	0.088
(Adults only) Systemic Elimination Half LIfe or Clearance/BW	0.136
(Children included) Systemic Elimination Half LIfe or Clearance/BW	0.171

B. Pharmacodynamic Steps

	Central Estimate Log(GSD)
Active Site Availability/General Systemic Availability	**0.084**
Non-Immune Physiological Parameter Change/Active Site Availability	**0.230**
Immune Physiological Parameter Change/Active Site Availability	**0.568**
Reversible Non-Immune Mild Functional Reserve Capacity--Change in Baseline Physiological Parameter Needed to Pass a Criterion of Abnormal Function	**0.452**
Non-Immune Moderate Reversible or Irreversible Functional Reserve Capacity	**0.202**
Non-Immune Severe and Irreversible Functional Reserve Capacity	**0.000**
Reversible Immune Functional Reserve Capacity	**0.510**

Example—Central Estimates of Summary Overall Log(GSD's) for Various Ingested Systemic Toxicants

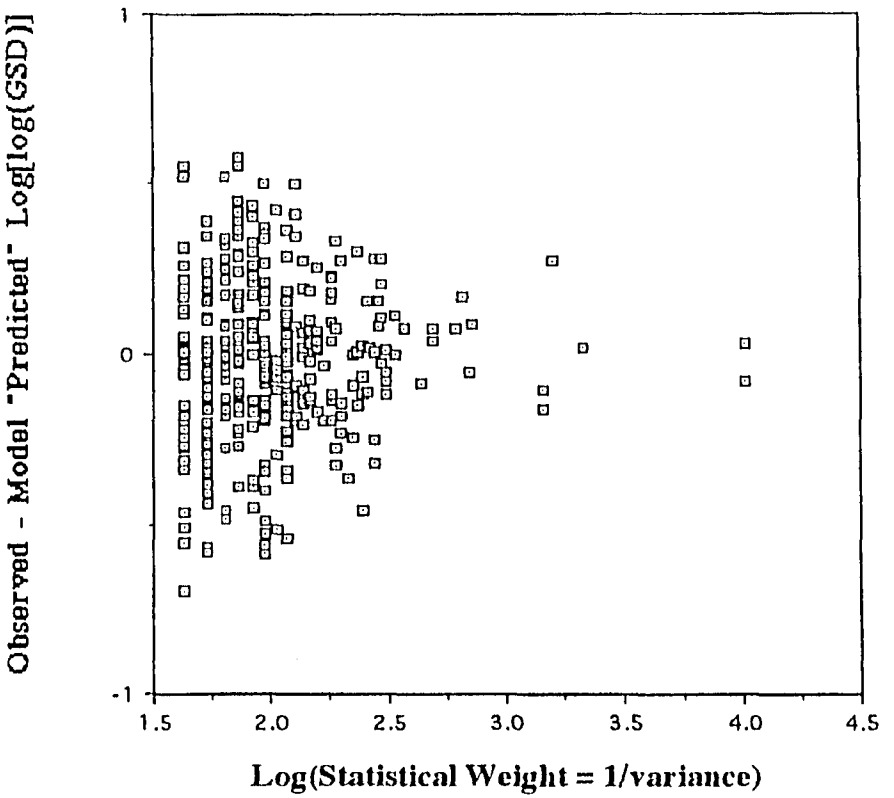
Route and Type of Response	Log(GSD)
Ingested Systemic Chronic Toxicant--mild reversible nonimmune effects	0.621
Systemic Chronic Toxicant--moderate reversible or irreversible nonimmune effects	0.471
Systemic Chronic Toxicant--severe irreversible effects	0.426
Chronic toxicity from an orally administered drug with perfect compliance (no contact rate variabity)--mild reversible non-immune effects	0.563
Chronic toxicity from an orally administered drug with perfect compliance--moderate reversible or irreversible non-immune effects	0.392
Chronic toxicity from an orally administered drug with perfect compliance--severe and irreversible non-immune effects	0.336
Acute toxicity from an orally administered drug with perfect compliance (no contact rate or elimination rate variability)--Mild reversible non-immune effects	0.536
Acute toxicity from an orally administered drug with perfect compliance --Moderate reversible or irreversible non-immune effects	0.352
Acute toxicity from an orally administered drug with perfect compliance--Severe irreversible non-immune effects	0.289

ASSESSING THE SPREAD OF VARIABILITY VALUES AMONG CHEMICALS AFTER CONTROL FOR TYPE OF TOXICITY AND ROUTE OF EXPOSURE

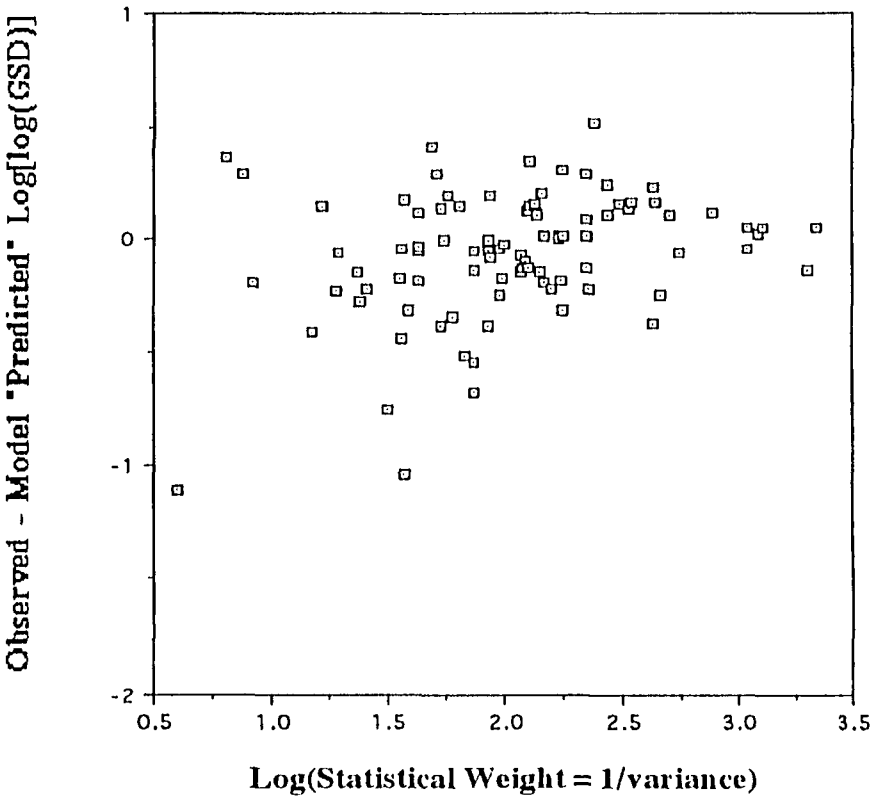
Funnel Plots--Show Tendency for Reduced Model Prediction Error for Stronger Data Points

Ideally, With Increasing Statistical Power, Measurement Error Becomes Small Relative to Real Variation Among Chemicals in Interindividual Variability in Susceptibility

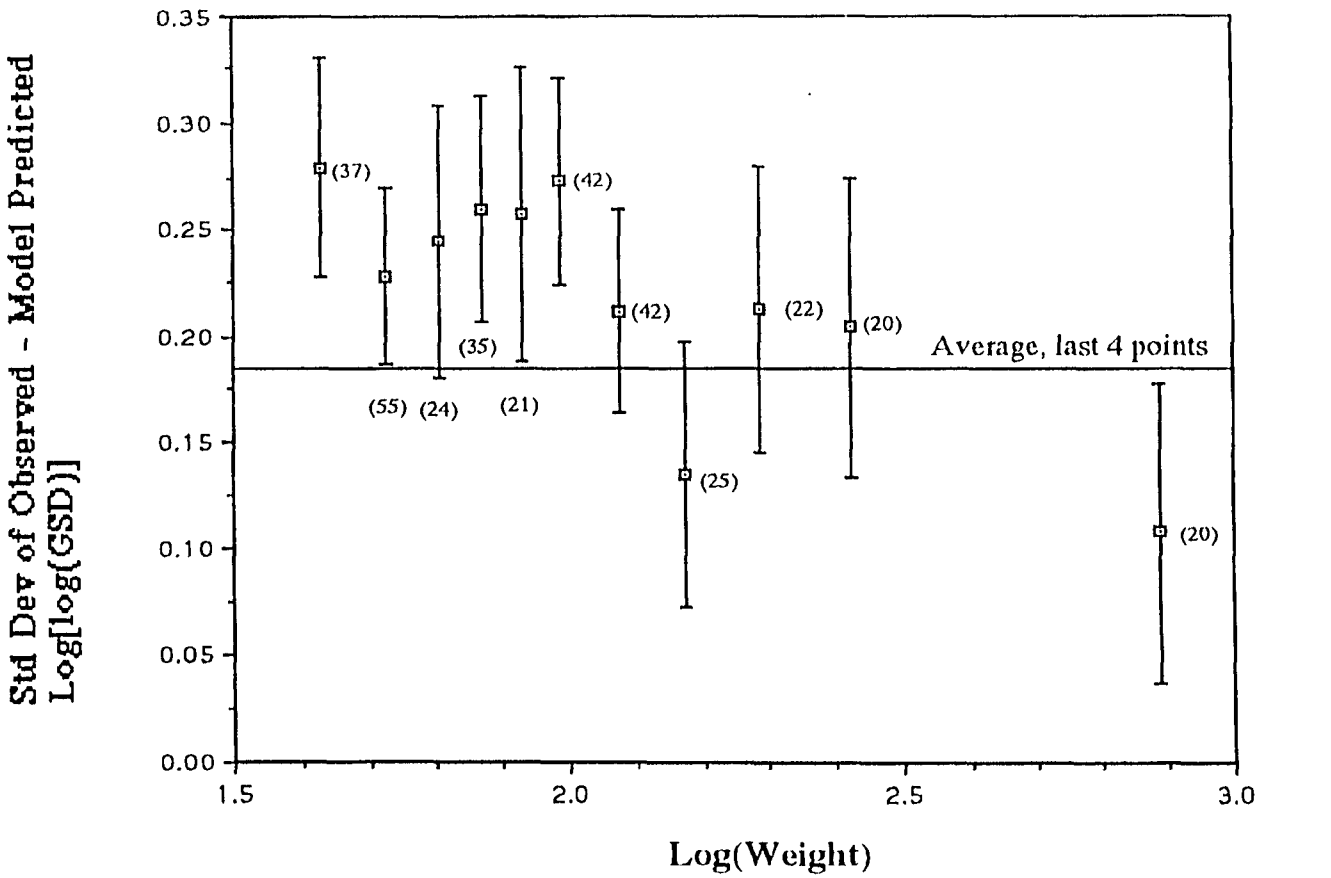
"Funnel Plot" for Pharmacokinetic Interindividual Variability Observations



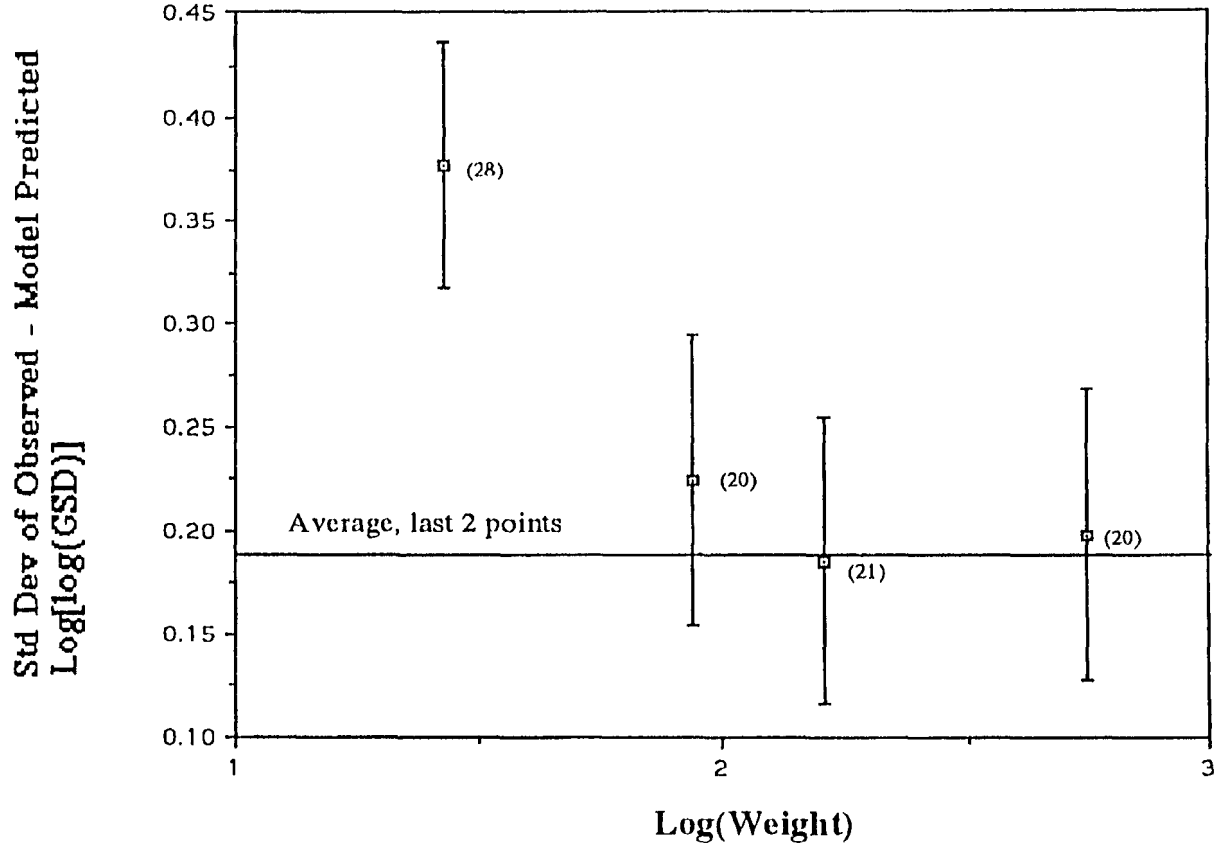
"Funnel Plot" for Pharmacodynamic Interindividual Variability Observations

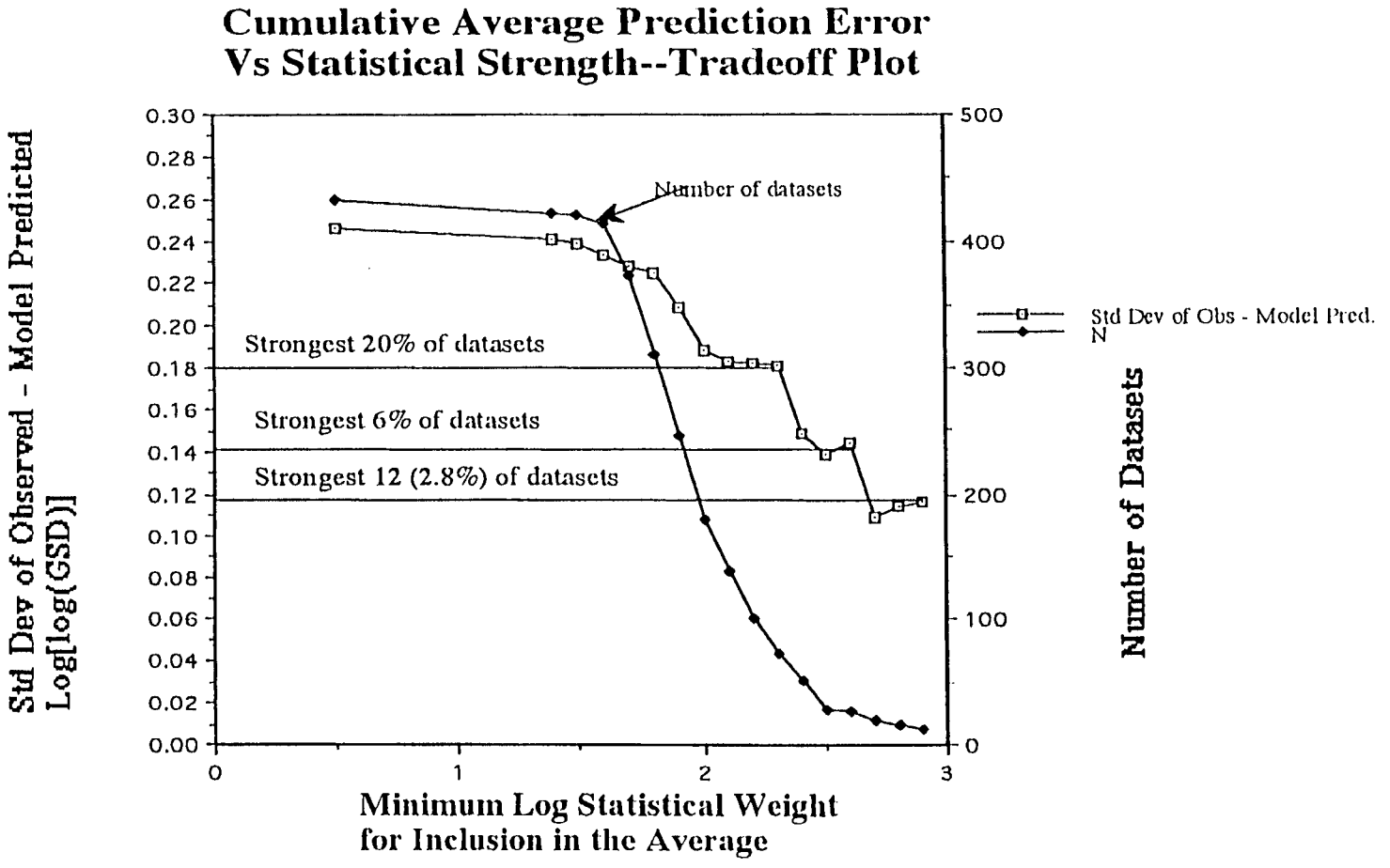


Prediction Error and Log(Statistical Weight) for Pharmacokinetic Interindividual Variability Observations



Prediction Error and Log(Statistical Weight) for Pharmacodynamic Interindividual Variability Observations

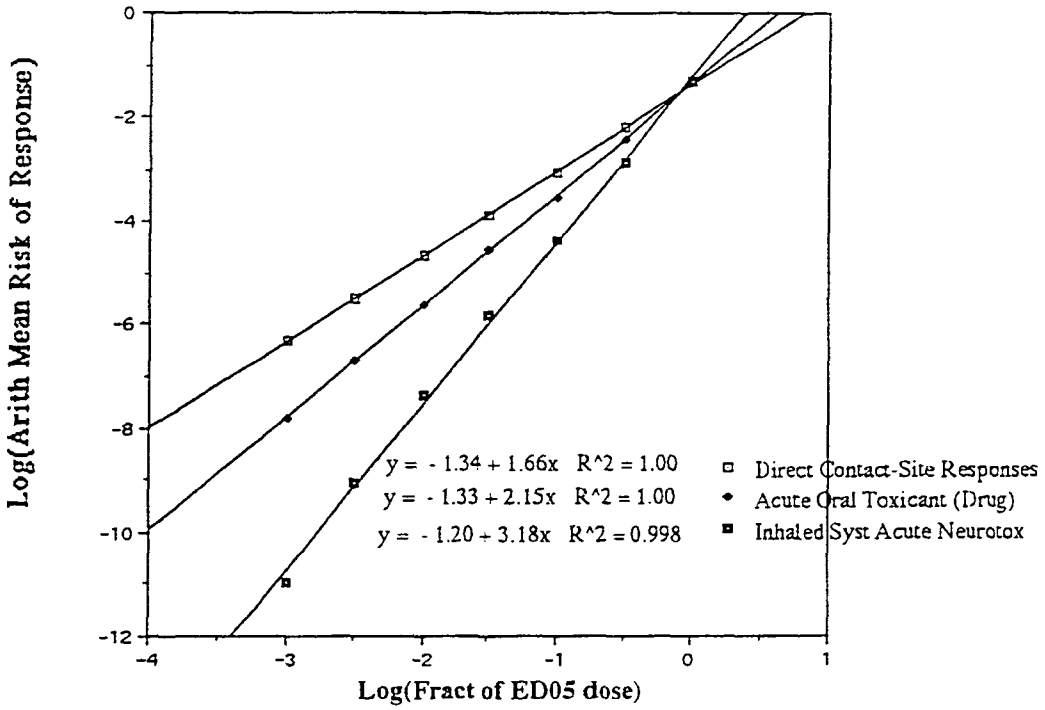



Cumulative Average Prediction Error
Vs Statistical Strength--Tradeoff Plot
Std Dev of Observed - Model Predicted
Log[log(GSD)]]
0.30
0.28
0.26
0.24
0.22
0.20
0.18
0.16
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00
500
400
300
200
100
0
Number of datasets
Strongest 20% of datasets
Strongest 6% of datasets
Strongest 12 (2.8%) of datasets
Std Dev of Obs - Model Pred.
N
Number of Datasets
0
1
2
3
Minimum Log Statistical Weight
for Inclusion in the Average

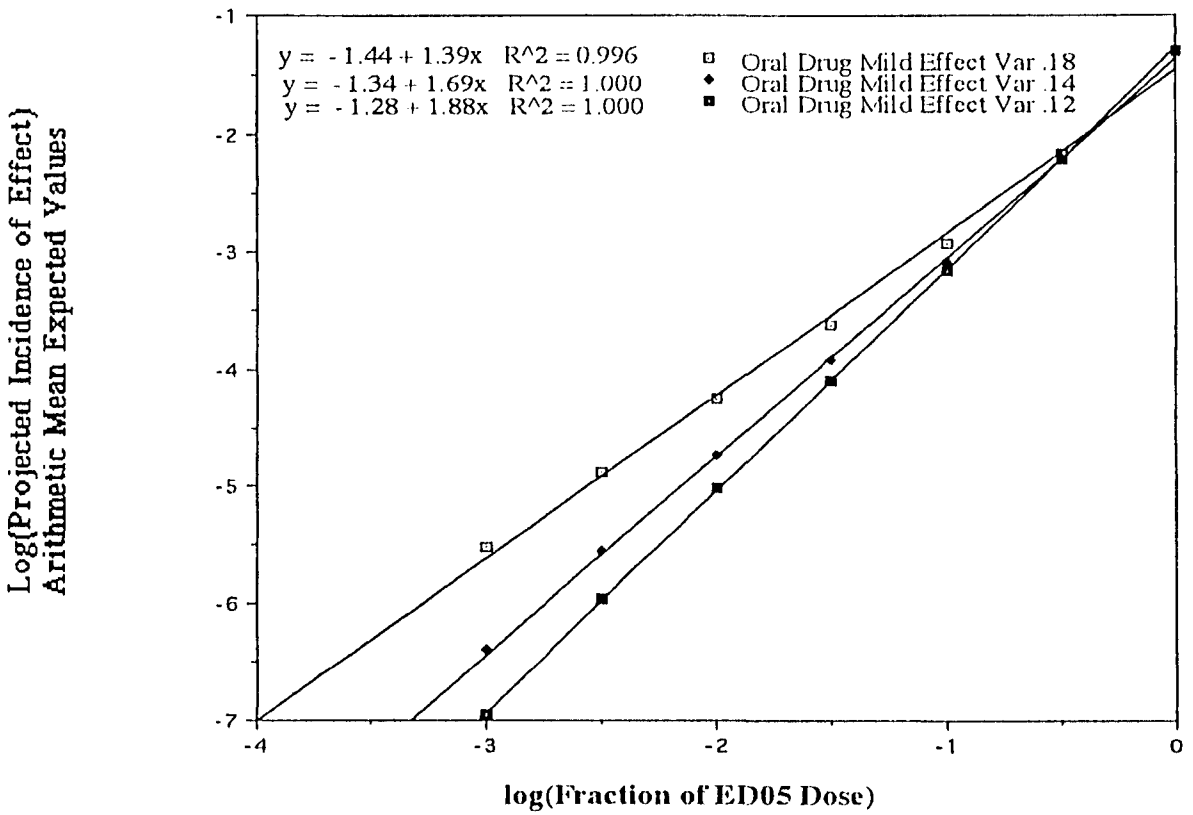
Incidence of Effects Projected At Various Fractions of a Human ED05 Dose

Type of Agent	Central Estimate of Log(GSD)	Std Dev of Log[log(GSD)]	Fraction of EDO5 1	.31	0.1	.032	.01
Orally administred drug, mild chronic effect	0.563	0.12	5.0E-02	6.3E-03	7.0E-04	8.0E-05	9.4E-06
		0.14	5.0E-02	6.5E-03	8.4E-04	1.2E-04	1.8E-05
		0.18	5.0E-02	7.0E-03	1.2E-03	2.4E-04	5.5E-05
Orally administered drug with severe chronic effect	0.336	0.12	5.0E-02	1.4E-03	3.8E-05	1.0E-06	2.1E-08
		0.14	5.0E-02	1.6E-03	6.2E-05	2.6E-06	9.8E-08
		0.18	5.0E-02	2.1E-03	1.4E-04	1.2E-05	1.1E-06
Orally administred drug, mild acute effect	0.536	0.12	5.0E-02	5.6E-03	5.6E-04	5.8E-05	6.1E-06
		0.14	5.0E-02	5.8E-03	6.9E-04	9.0E-05	1.2E-05
		0.18	5.0E-02	6.3E-03	1.0E-03	1.9E-04	4.1E-05
Orally administered drug with severe acute effect	0.289	0.12	5.0E-02	7.9E-04	1.2E-05	1.6E-07	1.4E-09
		0.14	5.0E-02	9.3E-04	2.2E-05	5.4E-07	9.7E-09
		0.18	5.0E-02	1.3E-03	6.4E-05	3.8E-06	2.0E-07

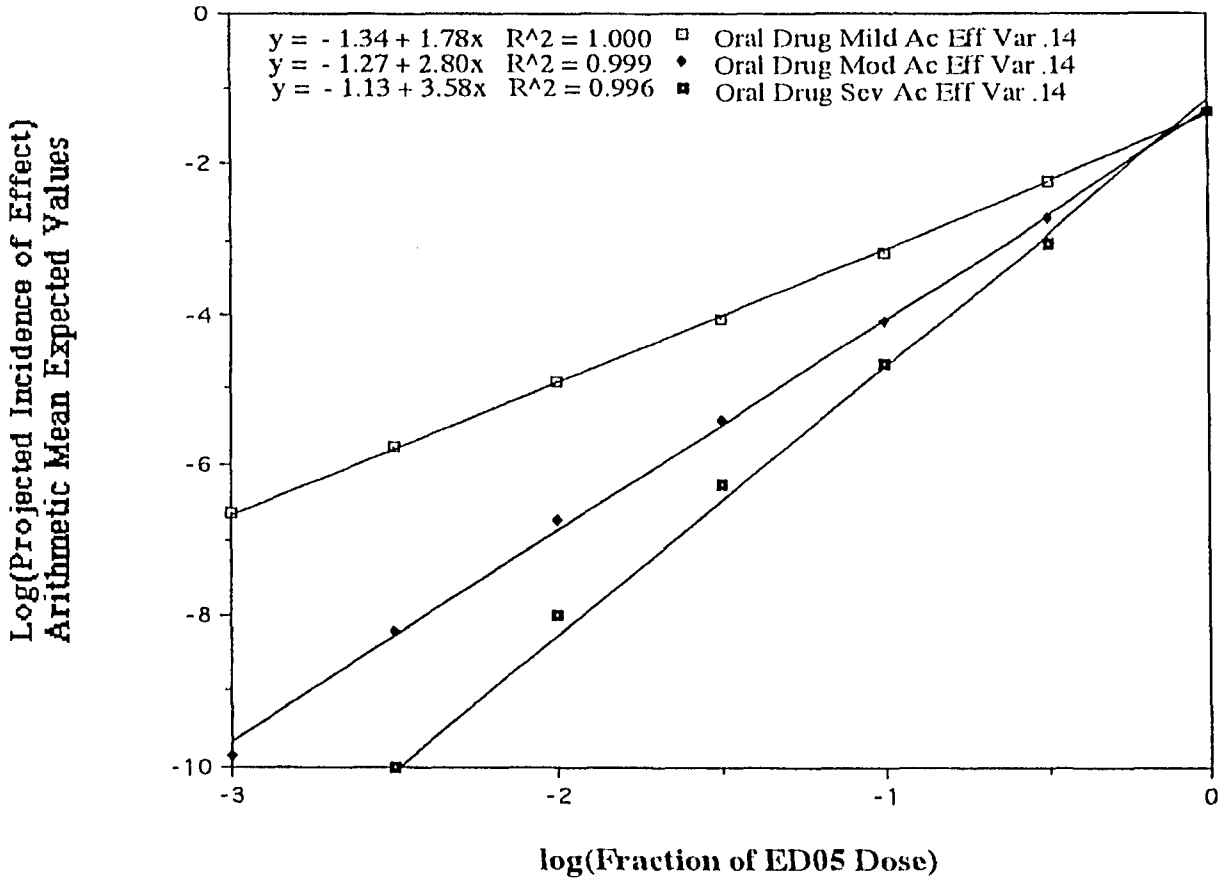
9/98 Version of the DataBase
Log Log Plots of Model Projections of the Mean Risk of Toxicant Exposures at Various Fractions of an ED05 Dose or Exposure Level



Power Law Plots Illustrating the Effect of Different Assumptions About the True Chemical-to-Chemical Variability of Log(GSD)'s --Chronic Toxicity for an Orally Administered Drug with Mild Effects



Power Law Plots Illustrating the Model-Estimated Implications of Different Degrees of Severity for Risk As a Function of Dose--Acute Toxicity for an Orally Administered Drug



Appendix I

Presentation Overheads

David W. Gaylor
Sciences International, Inc.

and

Ralph L. Kodell
National Center for Toxicological Research
U.S. Food and Drug Administration

Risk-based Reference Doses

David W. Gaylor*
Sciences International, Inc.
Alexandria, VA 22314

and

Ralph L. Kodell
National Center for Toxicological Research
U.S. Food and Drug Administration
Jefferson, AR 72079

David W. Gaylor, Ph.D.
Sciences International, Inc.
13815 Abinger Court
Little Rock, AR 72212
Ph: 501-228-9773
dgaylor@sciences.com

* This work was done when at the National Center for Toxicological Research.

Abstract

Reference doses (RfDs) for toxic substances are established to confine human exposures to only nontoxic or minimally toxic levels. Typically RfDs are calculated by dividing no observed adverse effect levels (NOAELs) or low observed adverse effect levels (LOAELs) by a series of uncertainty factors. Among these uncertainty factors is one for interindividual sensitivity, typically assigned a value of 10. If information is available on interindividual sensitivity, this default factor can be replaced with a factor expected to provide protection for a specified proportion of a population. To illustrate the procedure, examination of published databases suggests a standard deviation of the logarithm (base e) of individual sensitivity to be on the order of 1.7, *i.e.*, a factor of 5.5. Using this information in combination with an RfD based on a benchmark dose associated with a specified level of risk, the risk at the RfD can be estimated. For example, a benchmark dose associated with a risk of 10% divided by 60 is expected to limit risk at the RfD to about 1 in 10,000. This would replace an RfD having an unknown risk based on the LOAEL divided by 100.

Current Noncancer Safety Assessment

Based on NOAEL:

$$ADI\ (RfD) = NOAEL/U_A \bullet U_H \bullet U_S \bullet U_C \bullet M$$

U_A = uncertainty of animal to human extrapolation

U_H = sensitive individuals

U_S = extrapolation from subchronic data to chronic effects

U_C = additional sensitivity of children

M = modifying factor

Based on LOAEL:

UL = ratio of LOAEL to NOAEL

Default Values: $U_A = U_H = U_S = U_C = U_L = 10$

Worst Case: ADI (RfD) = LOAEL / 10,000 (EPA, 1991)

Benchmark Dose

Propose a benchmark dose estimated to produce an excess disease incidence of 10% (ED10).

With typical bioassays, this is near the lowest incidence that can be estimated with adequate precision and tends to be near the lowest observed adverse effect level (LOAEL).

Use a lower confidence limit on the benchmark dose (LED10) to account for experimental variation (EPA, 1996).

Confidence Associated with the Product of Uncertainty Factors

Consider the U's as independent random variables.

$$U = U_A \times U_H \times U_L \times U_S$$

$$\ell n U = \ell n U_A + \ell n U_H + \ell n U_L + \ell n U_S$$

Approximately normally distributed.

Products of default values of 10 provide approximately 99% coverage.

Statistical Distribution of U

$$\ell nU = \ell nU_A + \ell nU_H + \ell nU_S + \ell nU_L$$

Mean (median) of ℓnU:

$$\ell n\overline{U} = \ell n\overline{U}_A + \ell n\overline{U}_H + \ell n\overline{U}_S + \ell n\overline{U}_L$$

Standard deviation of ℓnU:

$$S_{\ell nU} = (S^2_{\ell nU_A} + S^2_{\ell nU_H} + S^2_{\ell nU_S} + S^2_{\ell nU_L})^{1/2}$$

Estimated Percentiles

$$\overline{U}_A = 1 \quad \ell n \overline{U}_A = 0$$

$$\overline{U}_H = 1 \quad \ell n \overline{U}_H = 0$$

Percentile for the normally distributed ℓnU is estimated by

$$\ell nU^* = (\ell n \overline{U}_S + \ell n \overline{U}_L) + Z s_{\ell nU}$$

$$U^* = \exp[(\ell n \overline{U}_S + \ell n \overline{U}_L) + Z S_{\ell nU}]$$

$$= \overline{U}_S \times \overline{U}_L \times \exp(Z S_{\ell nU})$$

$Z = 1.645$ for 95th percentile

$Z = 2.327$ for 99th percentile

Nonrandom Variables

If the RfD is based on human data:

$$U_A = 1 \text{ and } S_{\ell n U_A} = 0$$

If the RfD is based on the NOAEL:

$$U_L = 1 \text{ and } S_{\ell n U_L} = 0$$

If the RfD is based on chronic exposure data:

$$U_S = 1 \text{ and } S_{\ell n U_S} = 0$$

Intraspecies Variation (U_H)

U_H = Individual/median

U_H = 10 at 92nd percentile

U_H = 15 at 95th percentile

$S_{\ell nU_H} = 1.64$

Dourson and Stara (1983)

Table I. Estimated doses for specified levels of risk assuming a lognormal distribution with a standard deviation of 1.7 (log base e) for intraspecies variation.

Risk	Standard deviations below the median	Factor below the median	Factor below the BMD[a]
1 in 10	1.28	8.8	1.0
1 in 100	2.33	52.5	6
1 in 1000	3.09	191	22
1 in 10,000	3.72	558	63
1 in 100,000	4.26	1400	159
1 in 1,000,000	4.75	3210	365

Risk at the RfD(RfC) Based on the BMD_{10}

$BMD / (U_L = 10) \times (U_H = 10) = BMD / 100$

Estimated risk is 3 in 100,000

Assuming intraspecies variability standard deviation is 1.7 (log base e) ,i.e., a factor of 5.5

SUMMARY

The proposed method does not require extrapolating a dose response curve below an estimable benchmark dose.

Replaces uncertainty factors for LOAEL to NOAEL and intraspecies variation with a probabilistic approach.

Assumes a log normal distribution of intraspecies variability and requires an estimate of the standard deviation.

Provides estimates of the risk at the RfD(RfC) or at any specific exposure level.

Can calculate an RfD(RfC) with a specified level of risk.

Can obtain confidence limit by using the lower confidence limit on the BMD.

Can use in conjunction with uncertainty factors for extrapolation from animals to humans and from subchronic to chronic exposures.

Risk estimates can be improved with estimates of intraspecies variability for a specific chemical, class of chemicals, and/or specific biological effects.

References

1. Barnes, D.G. and M.L: Dourson. Reference dose (RfD): Description and use in health risk assessments. *Regulatory Toxicol. Pharmacol.* **8**: 471-486 (1988).
2. Gaylor, D.W. and Slikker, W., Jr. Risk assessment for neurotoxic effects. *NeuroToxicology* **11**: 211-218 (1990).
3. Dourson, M.L., S.P. Felter, and D. Robinson. Evolution of science-based uncertainty factors in noncancer risk assessment. *Regulatory Toxicol. Pharmacol.* **24**: 108-120 (1996).
4. Dourson, M.L. and J.F. Stara. Regulatory history and experimental support of uncertainty (safety) factors. *Regulatory Toxicol. Pharmacol.* **3**: 224-238 (1983).
5. Hattis, D. Strategies for assessing human variability in susceptibility and using variability to infer human risks. In: *Characterizing Human Variability in the Risk Assessment Process*, D. Neumann (ed.). ILSI Press, Washington, D.C. (1988).
6. International Programme on Chemical Safety. Environmental Health Criteria No. 170: Assessing Human Health Risks of chemicals: Derivation of Guidance Values for Health-Based Exposure Limits. World Health Organization, Geneva (1994).
7. Allen, B.C., Kavlock, R.J., Kimmel, C.A., and Faustman, E.M. Dose-response assessment for developmental toxicity. *Fundamental Appl. Toxicol.* **23**: 487-495 (1994).
8. Gaylor, D.W. and Kodell, R. L. Percentiles of the product of uncertainty factors for establishing probabilistic reference doses. *Risk Analysis* **20**: 245-250 (2000).

Appendix J

Presentation Overheads

Lynne Haber
Toxicology Excellence for Risk Assessment (TERA)

and

Michael Dourson
TERA

Use of Categorical Regression to Characterize Risk Above the RfD

Lynne Haber and

Michael Dourson

Toxicology Excellence for Risk Assessment (*TERA*)

Definitions of RfD

An RfD is an estimate (with uncertainty spanning perhaps an order of magnitude) of a daily exposure to the human population (including sensitive subgroups) that is likely to be without an appreciable risk of deleterious effects during a lifetime.

RfD Definition	Regression model
"is likely to be"	$P(*) > 0.95$
"without appreciable risk"	$r < 10^{-2}$ (??)
"deleterious effect" AEL	severity =

New RfD Definition

$P (r < 10^{-2} \text{ at dose} < RfD) > 0.95$

where $r = P$ (severity $\geq$ AEL)

Categorical Regression

- Toxicologist judgment
  - » Each dose assigned severity level

    0 = no effects observed

    1 = minimal effects

    2 = moderate - severe adverse effects

    3 = extreme or lethal effects
  - » Can include multiple studies, incidence, mean, or qualitative data
  - » Can evaluate separately (stratify) by endpoint, by species, etc.
- Mathematical analysis
- Results judged by data quality, statistics, graphics

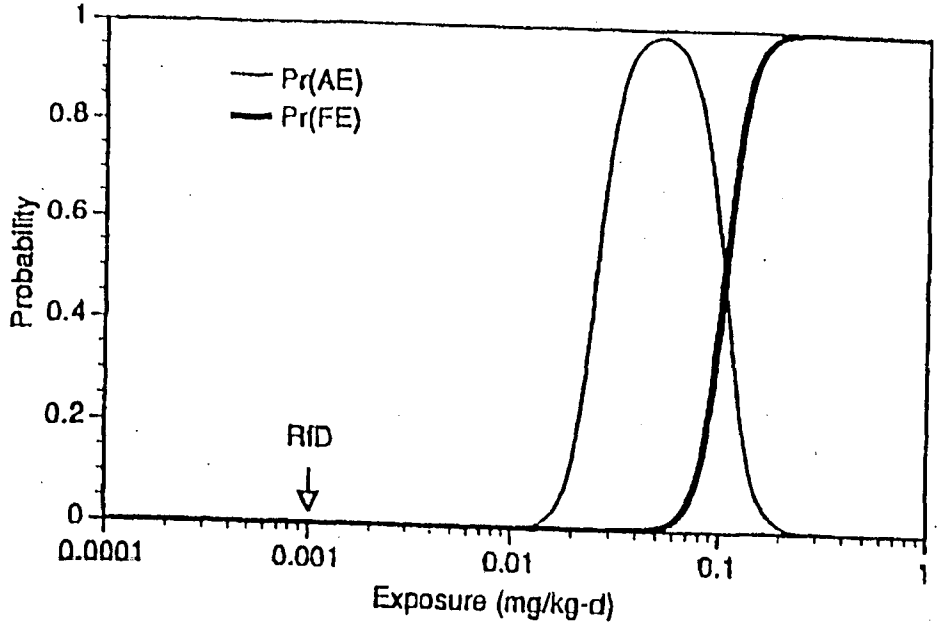
Advantages and Limitations of Categorical Regression

- Advantages
  - » All useful data can be categorized and included in quantitative analysis
  - » Can apply when data inadequate for calculating ED10
  - » Accounts for severity of effect
  - » Meta-analysis possible
  - » Can take duration into account
- Limitations
  - » Animal to human extrapolation
  - » Loss of information

Frequency of Categories of Effects Associated with Aldicarb Exposure in Humans

Dose	Group Size	Frequency of Responders within Category		
		NOAEL	AEL	FEL
0	22	22	0	0
0.01	8	8	0	0
0.025	12	8	4	0
0.025	4	0	4	0
0.050	12	1	11	0
0.050	4	0	4	0
0.075	4	0	4	0
0.10	4	0	2	2

Adapted from Dourson et al. (1997)


Pr(AE)
Pr(FE)
Probability
1
0.8
0.6
0.4
0.2
0
RfD
0.0001
0.001
0.01
0.1
1
Exposure (mg/kg-d)

Probability of an Adverse or Frank Effect - Aldicarb

Dose	P(AE or FE)	Upper 95% CL
0.001 (RfD)	**--**	**0.00001**
0.003	--	0.0007
0.01	0.0014	0.04
0.015	0.03	0.17
0.02	0.14	0.36
0.025	0.44	0.67
0.03	0.79	0.93

Adapted from Dourson et al. (1997)

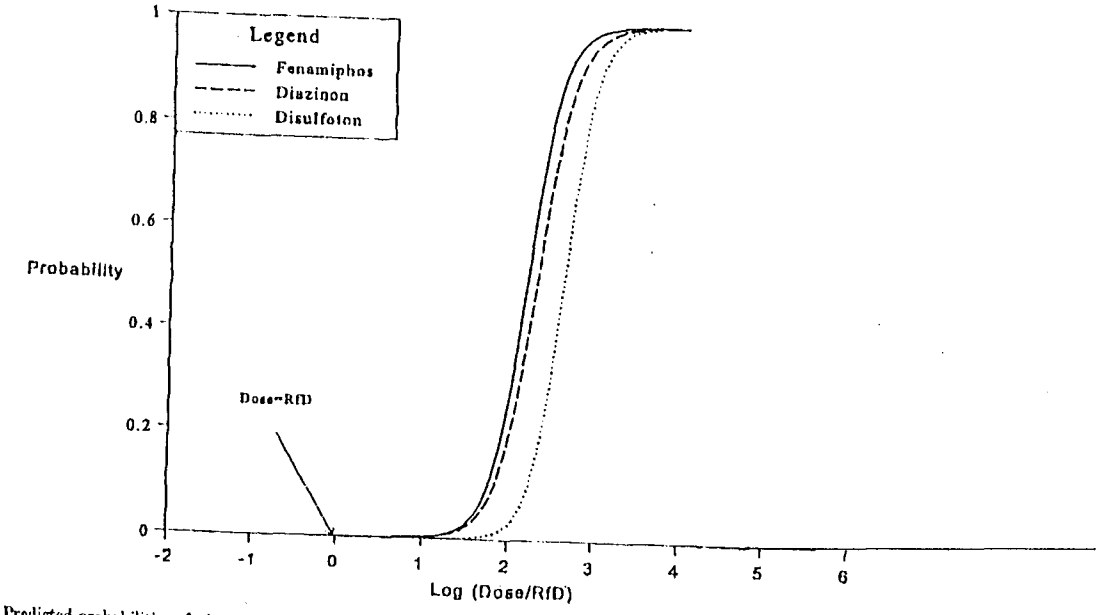
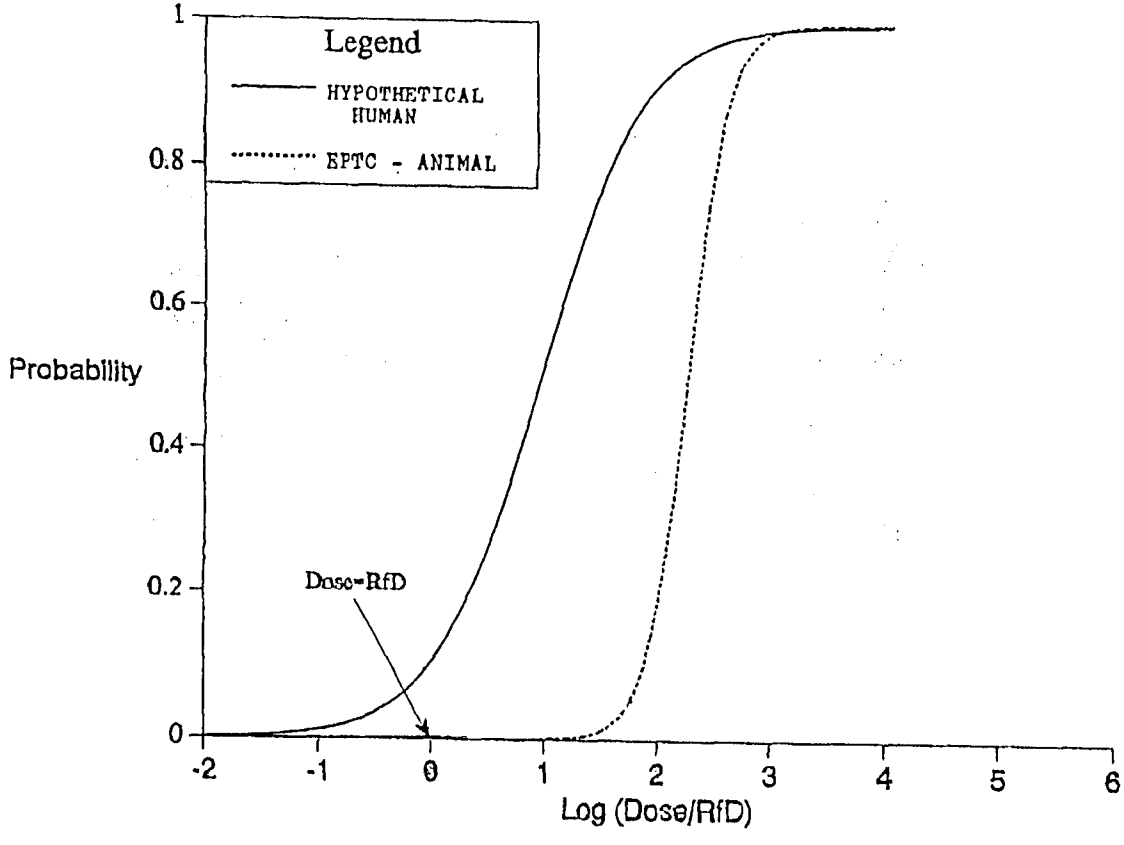
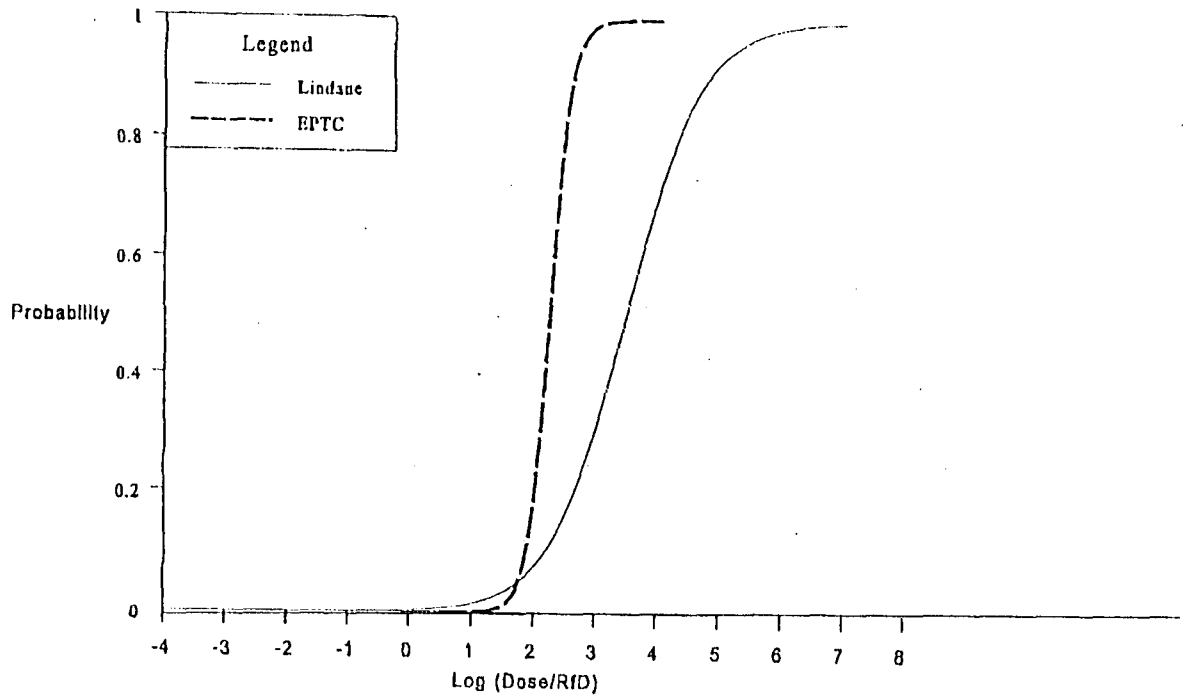


FIG. 1. Predicted probabilities of adverse or frank effects in humans after oral exposure to three pesticides. Three-category regression model. Doses scaled to human doses based on equivalence of (body weight)$^{3/4}$.


Legend
HYPOTHETICAL HUMAN
EPTC - ANIMAL
Probability
1
0.8
0.6
0.4
0.2
0
Dose=RfD
-2
-1
0
1
2
3
4
5
6
Log (Dose/RfD)

TEUSCHLER ET AL.



Issues

- Sensitive populations
- How account for use of UFs
- Model dependence
- Force model to go through 0 at RfD?
- Choice of data to model
- Rules for assigning severity categories
- Rules for combining studies
- Rules for model acceptance
- Policy regarding interpretation

Advantages of Categorical Regression

- All useful data can be categorized and included in quantitative analysis
- Can apply when data inadequate for calculating ED10
- Accounts for severity of effect
- Meta-analysis possible
- Can take duration into account
- Consistent basis for calculating risk above RfD

Uncertainty Factors

RfD = NOAEL/UF

Variability/Variation (CSAF)

- Interspecies -tk and td
- Intraspecies - tk and td

Uncertainty

- LOAEL/NOAEL
- Subchronic to chronic
- Database

Issues for Use of Other Approaches

- Not all UFs are equal
- Year of assessment affects what a UF of 10 means
  - » Initially - default
  - » Now - judgement that data are insufficient to reduce
  - » Future - may be CSAF
- Distributions for UFs need to reflect data supporting UF
- What use for distributions when use CSAFs?

Appendix K

Presentation Overheads

Reisha Putzrath
Georgetown Risk Group

We should make our models as simple as possible, but no simpler.

You can't get out of a problem by using the same thinking it took to create the problem.

– Albert Einstein

What Is the Issue to Be Resolved?

- Replace the RfD/RfC method and assumptions, e.g., with distributional analyses

- Use RfD/RfC, but expand method to evaluate exposures above these levels

- Make the current (or alternative) method more amenable to combine with other information

Assume Current Methods, but Estimate Risks:

- Above the RfD and
  - between RfD/RfC and NOAEL or LED_x
  - between NOAEL or LED_x and LOAEL or ED_x

- Carcinogens with curvilinear dose-response curves

- Margins of exposure?

Curvilinear Dose-response Curves for Carcinogens

- Must have data to reject the default. Use these data to estimate dose-response curve and its upper bound.

- If have reason to believe dose-response curve is not the same at exposure of interest, *but no data*, use policy decisions rather than mathematics.

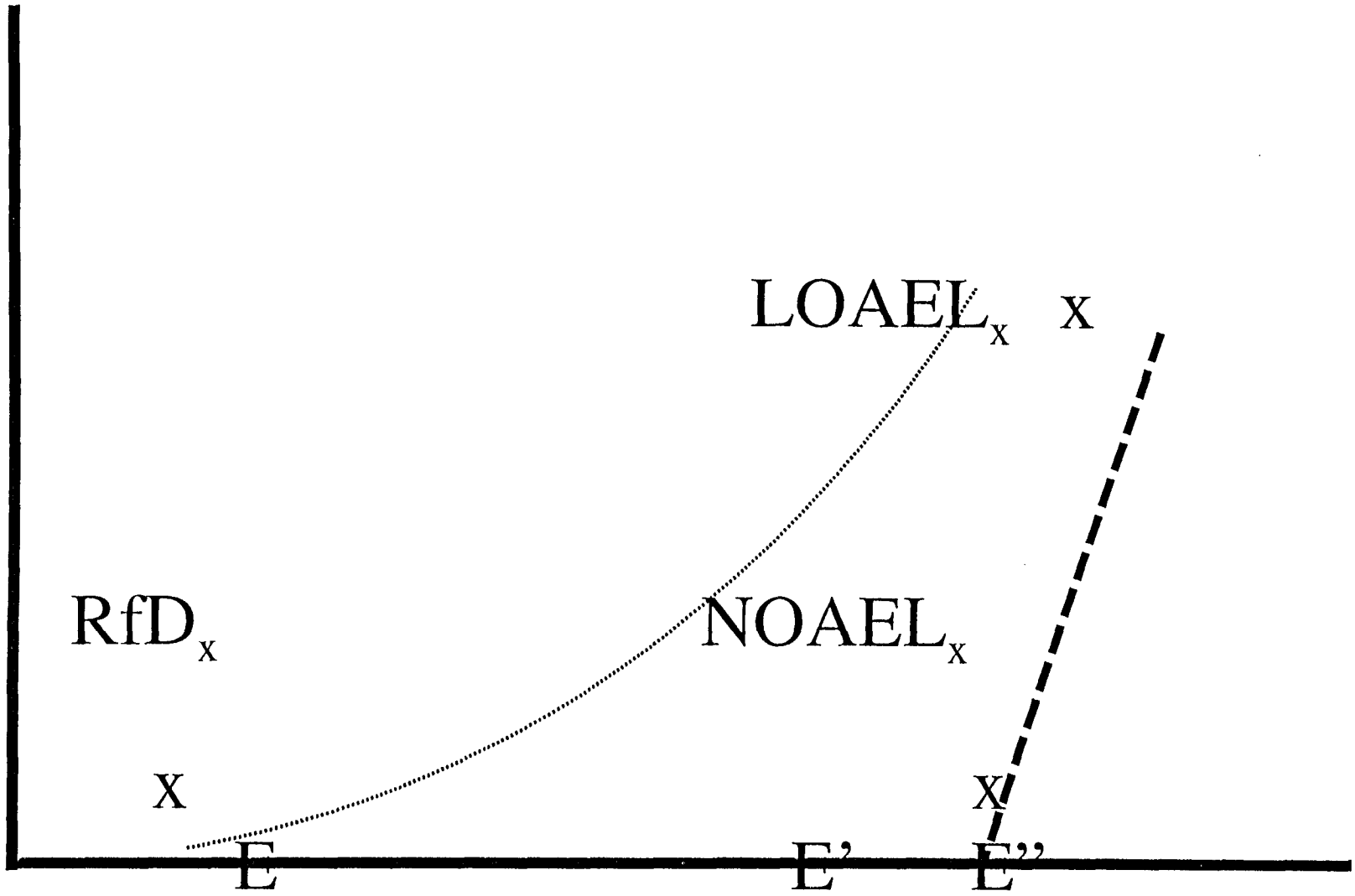
- The true dose-response curve and its upper bound can not both be continuous functions through the origin (Putzrath 2000).

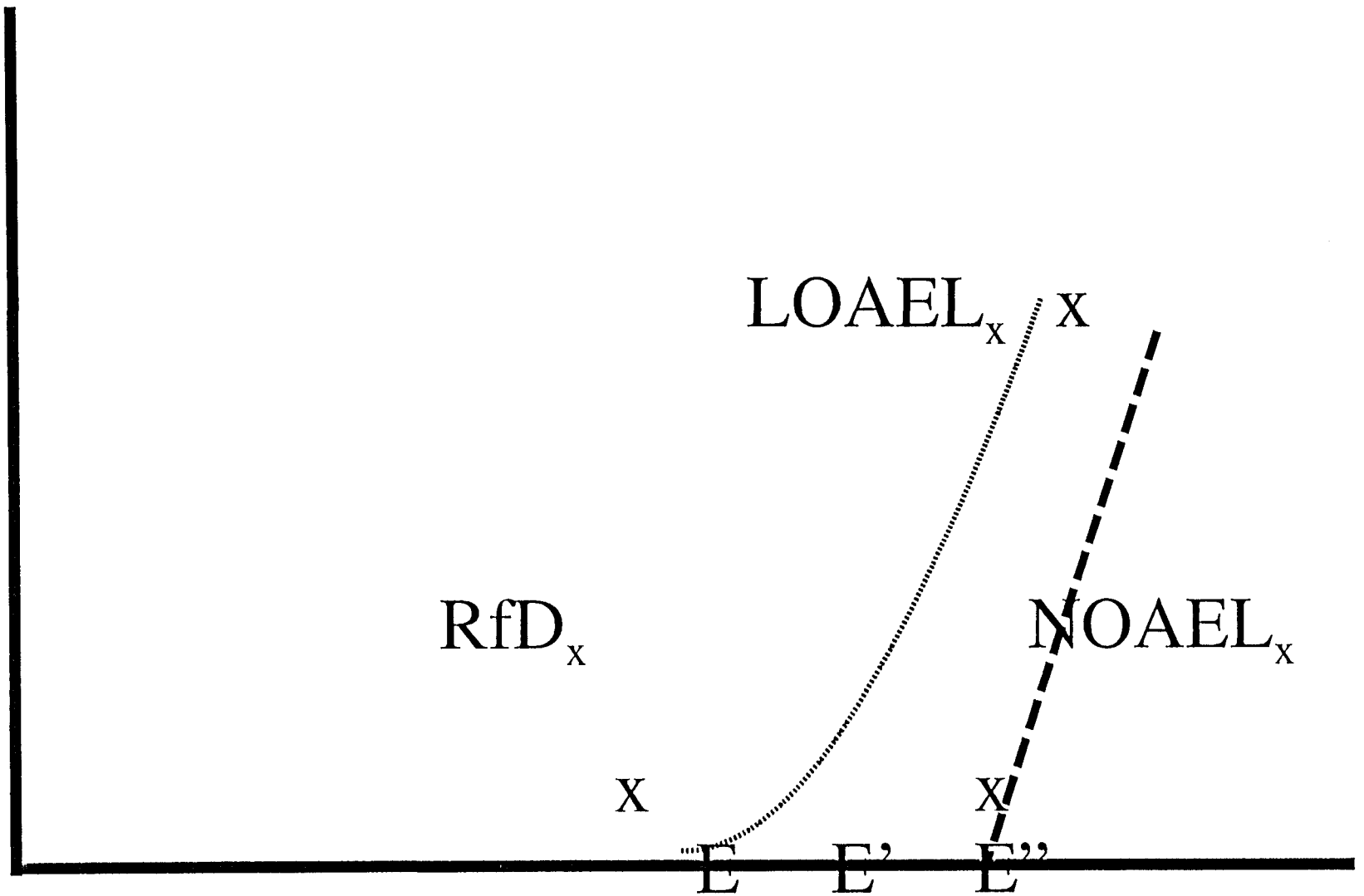
Use a Tiered Approach, as Recommended by NAS

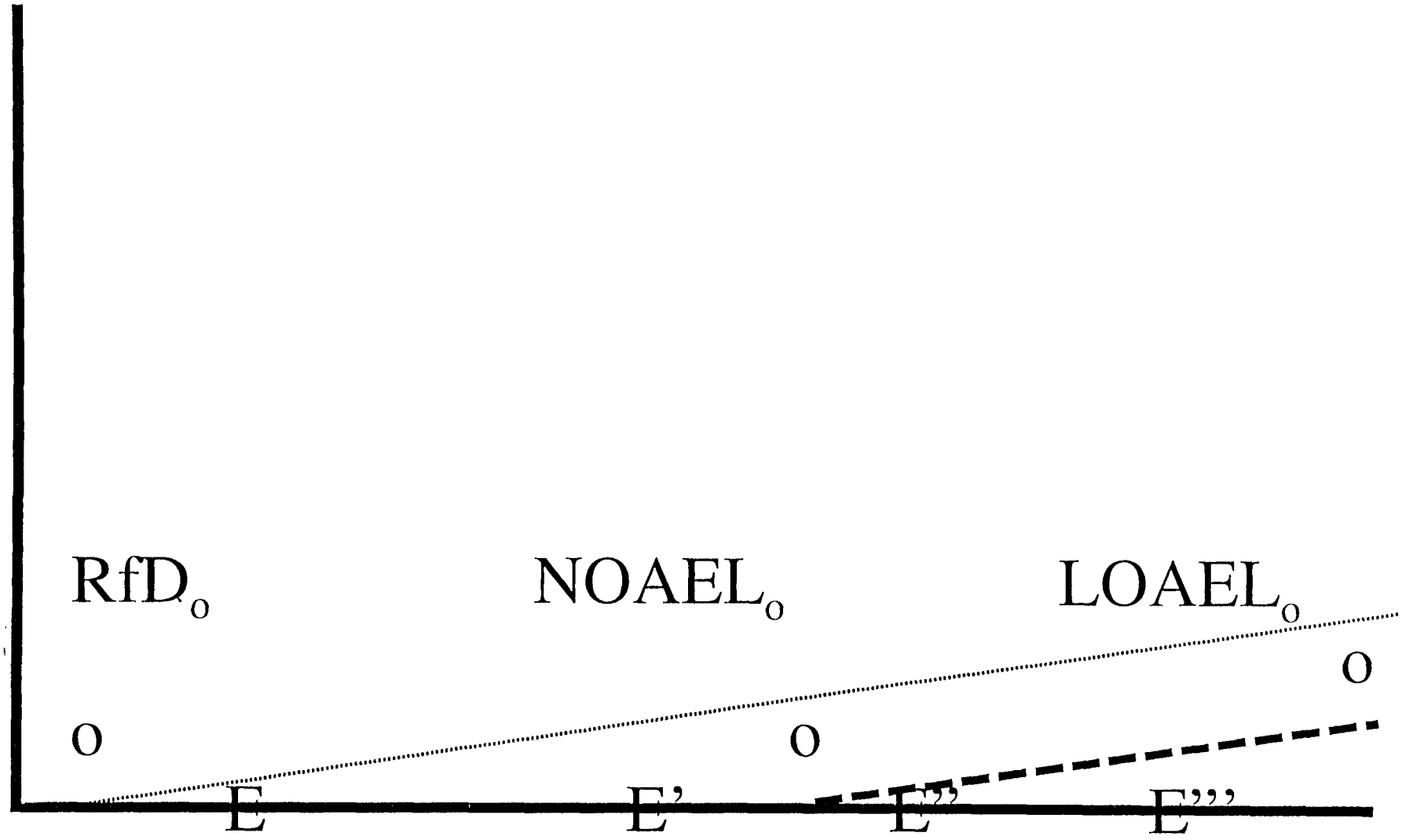
- Tier between bright line and distributional analyses: look at RfD/RfC, NOAEL, LOAEL, and exposure
- Generic distributions
- Distributions for similar chemicals

Characterize the Low-dose, Dose-response Curve

- Is the dose-response curve shallow or steep?
- How close to the NOAEL is the RfD?
- Is the exposure closer to the RfD, the NOAEL, or the LOAEL, i.e., how accurately must the curve and risk be estimated?


LOAEL$_x$
x
RfD$_x$
NOAEL$_x$
x
x
E
E'
E''


LOAEL$_x$
x
RfD$_x$
NOAEL$_x$
x
x
E
E'
E''


RfD$_o$
NOAEL$_o$
LOAEL$_o$
O
O
O
E
E'
E''
E'''

Uncertainty Factors Differ:

- *Human variability*: Distributions can characterize variability better than point estimates.

- *Interspecies*: Distribution or biologically based model?

- *LOAEL to NOAEL*: Distribution or chemical-specific estimate of dose-response curve?

- *Missing or poor quality data*: use surrogates or ?????

Thresholds

- Policy for non-cancer endpoints is that one exists.
  - Different distributions for cancer and non-cancer?
  - Which distributions have limits?

- How should the threshold be estimated?

- Will alternative methods eliminate thresholds, and what are the implications?

Appendix L

Presentation Overheads

Kenny Crump
ICF Consulting

Equations produced by Kenny Crump:

Use of lognormal distribution for extrapolating from higher to lower dose

$$P(e) = N\left(\frac{(Lne - a)}{\bullet}\right)$$

e = exposure
• = Ln (standard deviation)
a = set so that risk at BMD = 5 percent
N = normal distribution

Transfer function

eT do

$$P(e) = Pr(eT > do)$$

$$= Pr(e > do/T)$$

eT = internal dose
T = transfer relating external dose to internal dose (determined from pharmokinetic data)
Do = threshold dose (internal exposure) (determined from pharmacodynamic data)