United States
Environmental Protection
Agency

Office of Research and
Development
Washington DC 20460

EPA/630/R-96/010
September 1996

EPA

# Summary Report for the Workshop on Monte Carlo Analysis

RISK ASSESSMENT FORUM

EPA/630/R-96/010
September 1996

# SUMMARY REPORT FOR THE WORKSHOP ON MONTE CARLO ANALYSIS

U.S. Environmental Protection Agency
New York, NY
May 14-16, 1996

Risk Assessment Forum
U.S. Environmental Protection Agency
Washington, D.C. 20460

## NOTICE

The statements in this report reflect the views and opinions of the workshop panelists. They do not represent analyses or positions of the Risk Assessment Forum or the U.S. Environmental Protection Agency (EPA).

This report was prepared by Eastern Research Group, Inc. (ERG), an EPA contractor, and Menzie•Cura & Associates, Inc., a subcontractor to ERG, as a general record of discussion held during the Workshop on Monte Carlo Analysis (May 14-16, 1996). As requested by EPA, this report captures the main points and highlights of the meeting. It is not a complete record of all details discussed, nor does it embellish, interpret, or enlarge upon matters that were incomplete or unclear.

# CONTENTS

# SECTION ONE

# INTRODUCTION AND SUMMARY OF OPENING REMARKS

## 1.1 BACKGROUND AND PURPOSE

The importance of adequately characterizing uncertainty and variability in human health risk assessments has been emphasized in several U.S. Environmental Protection Agency (EPA) documents and activities. These include:

- The EPA Risk Assessment Guidelines.
- The Risk Assessment Council (RAC) Guidance or "Habicht Memo."
- The 1995 Policy for Risk Characterization.

There are several approaches to characterizing uncertainty and variability; however, Monte Carlo analysis is the most frequently encountered. EPA's Risk Assessment Forum has therefore undertaken the development of preliminary guidance on using Monte Carlo analysis.

Proponents of Monte Carlo analysis have criticized deterministic approaches to risk assessment, which results in point estimates of exposure and risk. They argue that point estimates are problematic because there is no way of knowing the degree of conservatism. Also, point estimates are sometimes derived from combinations of exposure factors that may be unrealistic. The proponents argue that Monte Carlo analysis and other quantitative uncertainty analysis techniques are superior to point estimates because they provide a full description of exposure where the impact of assumptions is explicit.

Critics of the application of Monte Carlo analysis techniques to risk assessment express concerns over the potential for misuse, particularly in data poor situations. They argue that a false sense of certainty or completeness is often associated with the results of Monte Carlo analyses and this could mislead risk managers. Also, critics are concerned that the value added

however, discussions were further stimulated by a presentation from panelist Max Henrion (Lumina Decision Systems, Inc.) on presenting information on uncertainty analysis.

Within each topic area, as the presentation concluded, the workshop chair facilitated a lively open discussion of the issues by panelists. Toward the end of a discussion session, observers were invited to comment and ask questions of the panel members.

On the evening of the second day (May 15), panelists were assigned to workgroups. Each workgroup was given a writing assignment to organize and clarify major points raised on a particular topic during the first two days of discussion. On the third day of the workshop (May 16), the workgroups presented the results of their writing assignments and panelists held final discussions to further clarify the emerging principles.

Section Two of this workshop report summarizes the discussions held during the workshop. Section Three provides the guiding principles and conclusions developed by the panelists. For issues where the workshop panel members were in general agreement, the principle is stated with supporting arguments. For those issues where the panel was divided, the pros and cons of various proposed principles are provided. The principles articulated in this report will provide the foundation for the development of future Agency policies and guidance related to Monte Carlo analysis. In the interim, this workshop report may provide useful perspective and insight to assist risk assessors in reviewing and preparing exposure assessments that use Monte Carlo analysis to characterize uncertainty or variability.

## 1.3 WELCOME AND REGIONAL PERSPECTIVE

Mr. William J. Muszynski, Deputy Regional Administrator, U.S. EPA Region 2

Mr. Muszynski opened the workshop by welcoming the participants and observers to EPA's Region 2 and to the Workshop on Monte Carlo Analysis. In giving his perspective on the importance of the workshop, Mr. Muszynski noted that efforts such as this are critical both to EPA's ongoing commitment to use the best possible science and to EPA's goal of better explaining how the Agency uses science. In fact, he said, using strong science and data is one of

EPA's seven guiding principles—and one that Administrator Browner frequently highlights when discussing the need to incorporate quality science into decision-making. Mr. Muszynski commented that holding a workshop on scientific issues related to Monte Carlo analysis will help expand recognition of EPA as a leader in environmental *science* as well as environmental *protection*.

To provide further context, Mr. Muszynski briefly summarized the history of the use of Monte Carlo analysis. He cited a textbook definition that described Monte Carlo analysis as a class of mathematical methods first used by scientists working on the development of nuclear weapons in Los Alamos in the 1940s. But use of Monte Carlo methods predates even the use of computers, he said. Indeed, one of the earliest documented uses of random sampling occurred in the late 1700s, when French scientists used the technique to define a solution to an integral. Scientists elsewhere used the technique to solve integral problems through the early 1900s, and Enrico Fermi performed Monte Carlo analysis calculations in the 1930s to study the behavior of the newly discovered neutron. More recently, use of Monte Carlo analysis in risk assessment—the subject of this workshop—has received growing attention. The fact that the workshop attracted such well qualified experts and such a range of observers, Mr. Muszynski said, demonstrates that interest in Monte Carlo analysis is high.

Mr. Muszynski concluded his remarks by enumerating the three major technical issues to be addressed in the workshop: input data/distributions for model parameters, evaluating variability and uncertainty, and presenting results. As a Deputy Regional Administrator, Mr. Muszynski stated that he was especially interested in the last topic, since he considers communicating variability and uncertainty to be the most difficult part of risk management.

Following this introductory statement, Mr. Muszynski asked Dr. William Wood to give an overview of the Risk Assessment Forum's goals for the meeting.

## 1.4 OVERVIEW

Dr. William Wood, U.S. EPA Risk Assessment Forum

Dr. Wood began by thanking Marian Olsen for organizing this Risk Assessment Forum-sponsored, region-based workshop. He then offered a brief explanation of the history and purpose of EPA's Risk Assessment Forum. Formed in the early 1980s in response to a National Academy of Sciences (NAS) study, the Risk Assessment Forum works to build consensus within the Agency on difficult and precedent-setting risk assessment issues. It does so by creating mechanisms (e.g., workshops) for dialogue and consensus-building. The ultimate products of this work include guidance documents and technical reports to communicate and promote consistency in the application of risk assessment methodologies. Recently issued drafts and publications include cancer, neurotoxicity, exposure, and ecological risk assessment guidelines.

Dr. Wood went on to describe the Risk Assessment Forum's membership, which includes 32 scientists from EPA laboratories, regions, and program offices. The members are distributed into four groups that address cancer, noncancer, exposure, and ecological assessment issues. The Exposure Oversight Group organized this meeting and will be responsible for followup. The Risk Assessment Forum typically uses a bottom-up process to develop documents. Technical panels work on individual issues and draft documents, and these are passed onto the oversight groups, the Risk Assessment Forum as a whole (which addresses technical issues), and finally the Science Policy Council (which addresses policy issues). This being a Risk Assessment Forum workshop, meeting participants will discuss technical rather than policy issues related to Monte Carlo analysis.

Agreeing with Mr. Muszynski that quantitative uncertainty analysis (including Monte Carlo analysis) has a long history, Dr. Wood focused on the history of Monte Carlo analysis within EPA. In 1986, EPA issued its first exposure assessment guidelines in which there was a discussion of uncertainty analysis. The Agency training courses that followed included a module that discussed Monte Carlo analysis. Nevertheless, EPA has in the past used Monte Carlo analysis mainly on an ad hoc basis, with those feeling comfortable with the tool using it and others largely ignoring it. To date, Monte Carlo analysis has been used mainly in exposure assessments. Its use in dose-response assessments is being investigated.

Most EPA programs are now experimenting with Monte Carlo analysis as a tool for quantitative uncertainty/variability analysis. Several forces are speeding the pace of this experimentation, including:

- The 1992 revised exposure assessment guidelines' call for fuller descriptions of exposure, which some have interpreted as encouraging the use of Monte Carlo analysis (which is, in fact, discussed in the guidelines).
- Strong recommendations from EPA's Science Advisory Board (SAB) to perform quantitative uncertainty analyses for risk assessments under SAB review.
- Recommendations from the National Academy of Sciences (NAS) on the use of quantitative uncertainty analysis to improve risk assessment (in NAS' 1994 report, *Science and Judgment in Risk Assessment*).
- Administrator Browner's 1995 risk characterization policy, which has provoked many discussions of quantitative uncertainty analysis.
- Risk-based legislation requiring EPA to discuss and clarify uncertainty in Agency risk assessments.

Dr. Wood noted that EPA faces both institutional and technical hurdles in attempting to broaden use of Monte Carlo analysis. These include a lack of EPA policy guidance on the use of Monte Carlo analysis, a lack of technical guidance, a lack of good examples of Monte Carlo analyses, limited experience and technical expertise with Monte Carlo analysis, a lack of guidance on how to distinguish between good and bad quantitative uncertainty analyses, a lack of default distributions to use in Monte Carlo analyses, concern that Monte Carlo methods are particularly subject to the "garbage in-gold out" phenomenon, concern about whether use of Monte Carlo analysis actually adds value to risk assessments, concern that the ability to model uncertainty using Monte Carlo techniques will be used as an excuse to not collect data, and concern about resource implications (given that reviewing Monte Carlo analyses takes much more effort than reviewing typical point estimate calculations).

Dr. Wood noted that this workshop is intended to address some of these hurdles. The Risk Assessment Forum hopes that the workshop will result in recommendations that the Forum can take to a technical panel. In the short term (over the next few months), the Risk Assessment Forum would then develop a set of guiding principles on Monte Carlo analysis.

These, in turn, would form the basis for a Science Policy Council policy statement on the use of Monte Carlo analysis as well as a more extensive guidance document on its use in EPA risk assessment. EPA also plans to develop a training course on Monte Carlo analysis.

After these remarks, Dr. Wood entertained questions from workshop observers. Two observers commented on widespread confusion about the use of terms related to uncertainty analysis (e.g., "uncertainty" versus "variability") and asked whether the Risk Assessment Forum plans to address these issues. Dr. Wood replied in the affirmative, stating that these questions will be addressed during and after the workshop. Another observer noted that the Risk Assessment Forum historically has addressed ecology rather than economics and asked whether the Forum plans to broaden its focus to consider economics. Dr. Wood stated that the Risk Assessment Forum has not addressed economics issues, but that the Science Policy Council is exploring that area. Several observers commented that avoiding policy issues during this workshop might be difficult because some questions (e.g., what is the scope of the assessment) straddle the border of science and policy. Dr. Wood agreed. Finally, several observers identified guidance material development efforts and other activities that have been or are being undertaken in their regions; they suggested that interested individuals seek out information on these activities.

Following this discussion, Dr. Wood thanked the participants and observers for attending the workshop and introduced Charlie Menzie, the workshop chair.

# SECTION TWO

# SUMMARY OF PANEL DISCUSSIONS

The agenda for the first and second days of the workshop was divided into three broad topic areas:

- Input data/Distributions for model parameters
- Variability/Uncertainty
- Presenting results

Each topic area commenced with one or more presentations to set the stage for subsequent discussion on the topic. The discussions, in turn, provided the foundation for crafting of principles and conclusions that are described in Section Three of this report. Panelist discussions are summarized below by topic area. Overheads and background papers supporting the presentations are provided in Appendix D.

## 2.1 DERIVING AND USING INPUT DATA AND DISTRIBUTIONS FOR MONTE CARLO ANALYSIS

This topic area opened with a presentation by David Burmaster on *Input Data/Distributions for Model Parameters.* Dr. Burmaster began by defining variability and uncertainty.

- **Variability** represents the natural heterogeneity or diversity in a well-characterized *population*. Variability is:
  - Usually not reducible through further measurement or study.
  - A *bounded* characteristic or property of the *population.*

- The primary physical, chemical, and biological phenomenon.

- **Uncertainty** represents ignorance (or lack of perfect knowledge) about poorly characterized phenomena or models. It is:
  - Sometimes reducible through further measurement or study.
  - An *unbounded* characteristic or property of the *analyst.*
  - The primary mental phenomenon.

He then briefly discussed the selection of key variables as random variables, as well as how processes create distributions. He also reviewed approaches to fitting univariate and bivariate distributions to data for variability; developing second-order random variables for uncertainty[1]; truncation of input variables; and correlations and/or dependencies among input variables.

Dr. Burmaster strongly recommended that, in any future EPA guidance documents about Monte Carlo methods and probabilistic assessment, the Agency include a statement at the beginning of the document to make it clear that 1) the report contains guidelines for minimum practices that are acceptable for use in probabilistic exposure assessments; 2) the report does not list all the possible techniques that a risk assessor may use for a particular assessment (since this would be impossible given the breadth and depth of probabilistic methods and their rapid development); 3) the Agency encourages the development and application of new methods in exposure assessments; and 4) the document should not be construed as limiting the development or application of new methods whose power and sophistication may exceed the guidelines for minimum acceptable practice contained in the document.

Overheads for David Burmaster's presentation are provided in Appendix D.

[1]Monte Carlo analyses in which variability or uncertainty is handled *separately* are referred to as two-dimensional ("2-D") or second-order analyses. The variables associated with uncertainty alone are sometimes referred to as second-order variables. Examples of these kinds of analyses can be found in Appendix D.

This was followed by presentations of two case studies illustrating how input data and distributions were derived for Monte Carlo analyses conducted as part of environmental risk assessments:

- The first case study (developed by Michael Dusetzina and delivered by Charles Menzie) covered the use of Monte Carlo analysis for a risk assessment for the benzene Maximum Achievable Control Technology (MACT) standard (under Title III of the Clean Air Act). The analysis was conducted as part of a screening-level risk assessment for 174 petroleum refineries. Dr. Menzie reviewed the purpose, scope, and overall methodology for the assessment. He also described which variables were used and how the variables and distributions were selected. He concluded by discussing the assumptions, uncertainties, and variabilities in the assessment. Overheads for the presentation and a paper describing this analysis are included in Appendix D.

- The second case study, presented by Dr. Teresa Bowers, concerned a risk assessment calculation done by Monte Carlo analysis for a Region V Superfund site. The assessment focussed on 10 contaminants of concern in a flood plain area. The analysis was a two-dimensional analysis[1] in which variability and uncertainty were "decoupled" and considered separately. Dr. Bowers reviewed the analytical methodology and described the approach and issues concerning selection of the concentration distribution and the exposure frequency distribution. A prepublication version of "The Use of Two-Stage Monte Carlo Simulation Techniques To Characterize Uncertainty and Variability" (Cohen et al.), describing this assessment, was distributed at the workshop; the paper will be published in the December 1996 issue of *Human & Ecological Risk Assessment: An International Journal.*

Workshop panelists then launched into a discussion of issues related to use of Monte Carlo analysis in general and derivation of input data/distributions for model parameters in particular, as summarized below.

### 2.1.1 Differences Between Deterministic Risk Assessment and Probabilistic Risk Assessment

Panelists spent some time comparing deterministic and probabilistic risk assessment. Probabilistic risk assessment not only provides distribution shapes, but also takes into account the dependency structure between the factors. Deterministic risk assessment does neither. These additional features of probabilistic risk assessment are important, one commenter said, because

many things with zero correlation have a strong dependence—for example, earned income as a function of age, and workplace exposures. These dependencies make a big difference in the calculation of any risk numbers.

Another important difference between deterministic and probabilistic risk assessment is that, in probabilistic risk assessment, the probabilistic expressions cannot be as easily inverted—there must be a deconvolution.

Another difference was that a deterministic risk assessment tends to focus on known data and ignore other influential factors. For example, a deterministic risk assessment for contaminated tap water might tend to focus on the ingestion exposure route, because this route can be quantified, and might tend to ignore the inhalation and dermal routes, which cannot be as easily quantified. A probabilistic risk assessment, on the other hand, ideally will at least consider, clearly acknowledge, and attempt to capture as many influential factors as possible.

Panelists briefly discussed the degree to which policy was embedded in a probabilistic risk assessment versus a deterministic risk assessment. They generally agreed that both approaches have a policy content because they both incorporate many assumptions. Also, how the results of a Monte Carlo analysis are used and what portion of the range is regulated will always remain a policy decision.

### 2.1.2 Use/Value-Added of Monte Carlo Analysis for Regulatory Decision-Making

A key topic of discussion was how the results of a Monte Carlo analysis are used for risk management decision-making and whether such an analysis added important value compared to a deterministic risk assessment. Panelists had varying impressions about the value of Monte Carlo analysis for decision-makers.

In one panelist's experience, Monte Carlo analysis appeared to be particularly useful when there were economic pressures on the risk management decision. Regions will need to understand what practical value is added by the sophistication of the Monte Carlo analysis. One

panelist mentioned that, to achieve consistency in how Monte Carlo analysis is applied to air risk management in California, state staff are developing distributions for use in these types of assessments. She encouraged EPA to develop distributions and require that they be used in Monte Carlo analysis assessments wherever that makes sense.

One panelist speculated that the interest of risk managers in probabilistic analysis likely is increasing, since EPA managers increasingly are called on by Congress to justify their risk management decisions.

Some panelists were skeptical about the appropriateness or value of using Monte Carlo results for risk management decision-making. Various concerns were expressed. One concern was that the results of a Monte Carlo analysis could be used to create the illusion of greater precision when in fact this was not the case. Another concern was over the difficulty of communicating the results of a Monte Carlo analysis to risk managers and the public. This theme was discussed further on the second day of the workshop—see Section 2.3.) One panelist strongly felt that Monte Carlo analysis is "overkill" for exposure analysis. Monte Carlo analysis was designed to solve complex equations and was not designed for simple multiple realizations of the world, he said. The panelist suggested that simple exposure models are sufficient for exploring the behavior of variance.

A number of panelists expressed a concern that risk managers would be confused by the plethora of results from a probabilistic analysis. One panelist stressed that it is important not to embark on a sophisticated analysis unless both analysts and managers know what they will do with the results.

Another panelist speculated that the results of probabilistic risk assessment could conceivably lead to reduced risk management costs. For example, when a full distribution of risk is available to risk managers, they may be able to establish more cost-effective cleanup levels. Incremental improvements in decision-making provided by probabilistic risk assessment versus deterministic risk assessment could sometimes make a big difference in the real world.

One panelist felt that microeconomics and cost-benefit analyses fit very nicely with probabilistic risk assessment, especially if they include the uncertainties as well as the variabilities. This leads to the discipline called the "value of information" where one can figure out what information is needed, how much society or an institution is willing to pay for it, and where to focus future research. For example, this discipline could address the question: Should an environmental agency spend money better defining body weights or investigating children's behavior at the playground? This type of question becomes answerable in a probabilistic framework.

Some panelists expressed concern about the validity of distributions—that without sufficient data to develop a distribution, there was a danger that a distribution was essentially a way of fabricating data and that such a distribution might not have any more validity than a point estimate. Another source of uncertainty is the fact that most biological processes are not well enough understood to be quantified. One panelist cautioned that probabilistic analysts must be very careful to neither have nor convey overconfidence in their results. Echoing this sentiment, a panelist pointed out that the results of a Monte Carlo become more uncertain toward the tails, yet it is these results that typically are used by regulatory agencies for decision-making. Even when decision-makers are told what the "error" is, will that help them make the decision, he asked?

One commenter pointed out that there are two types of uncertainty: uncertainty in measurement, and uncertainty in the assumptions underlying models—the cancer slope factors, for example. The latter type of uncertainty cannot be quantified and can only be dispelled by mechanistic research. This type of uncertainty makes Monte Carlo analysis irrelevant to risk assessment, he suggested, because it is substantial enough to make the results implausible. Other panelists disagreed. They felt this type of uncertainty could and should be modelled, where appropriate. Panelists agreed, however, that a Monte Carlo analysis should not be performed in situations where the results would be implausible.

It was suggested that guidance is needed on how to distinguish a good Monte Carlo analysis from a bad Monte Carlo analysis, so that decision-makers can tell the difference. Several panelists pointed out that the flaws in probabilistic risk assessment can be the same as in

deterministic risk assessment, but they are harder to see because they are cloaked in mathematics.

### 2.1.3 Use of a Tiered Approach/Steps in a Probabilistic Risk Assessment

Panelists strongly supported the idea of using a tiered approach to probabilistic analysis. For example, in the first tier analysts could define the range for each input and how important that was likely to be. In some cases that activity alone might be sufficient to meet the goals of the assessment. Several panelists stressed that the appropriate level of effort for an analysis depends on how the results will be used. An interactive educational dialogue with risk managers is important to define the appropriate level of analysis. Panelists then defined key steps in a probabilistic risk assessment.

#### *Problem and Purpose Definition*

One of the first steps in any assessment, panelists agreed, should be to clearly define the purpose of the assessment. This step clearly has policy dimensions. For example, one panelist recently was involved in a debate about whether to include pregnant women (who are largely in the 95th percentile) in a distribution or to create an entirely different exposure model with different distributions for the pregnant women. These types of questions have to be asked and answered long before beginning either a probabilistic or deterministic risk assessment, she stressed.

Supporting this idea, another panelist pointed out that how an analysis is constructed very much depends on the questions asked early on about the purpose of the analysis. He emphasized the importance of recognizing the philosophical underpinnings of the analysis—for example whether the goal is to specifically protect pregnant women, or to protect 95% of the population and basically assume that pregnant women are covered by the conservativeness of that approach.

A few panelists thought that many decision-makers currently lack the technical sophistication to frame the questions for a Monte Carlo analysis. However, another panelist said that, in her experience, risk managers typically look to the analyst for guidance about what is important. Risk managers do not need to be experts in probabilistic risk analysis.

A number of panelists supported the idea of bringing the various interested parties (e.g., risk managers, the public) in when formulating the problem and making the early philosophical or policy decisions. Doing this would help ensure that the analysis is responsive to the range of questions these parties might ask. Supporting this idea, a panelist pointed out that how a problem is framed has a large influence on how the assessment is formulated. At a hazardous waste site, for example, the analysis will be very different depending on whether the goal is to protect a specific population or to achieve a certain cleanup level. Regulatory agencies tend to have a predetermined sense of what endpoint they consider to be important, he said, but they should consult interested parties about their issues and concerns when formulating the analysis. They also should maintain flexibility so that they can respond to new concerns that might arise mid-stream.

Another panelist suggested that analysts should remain open-minded and be prepared to learn something during analysis—for example, about the question, the objectives, and how to refine the model from the failures (e.g., when the data do not quite fit).

In particular, panelists supported the idea of drawing an influence diagram when structuring a model. An influence diagram defines the relationships between what is currently known and what knowledge or data is desired. It should capture everything that is plausible. Then a conscious decision can be made about what factors can realistically be included in the model. In addition to its value for developing the model, an influence diagram provides a good mechanism for clearly communicating what was included or not included in the model. Panelists generally supported the idea of consulting or brainstorming with interested parties when possible about what should go in the influence diagram.

One panelist emphasized that, in deciding which variables to include, it is important to think about how the variables will be modelled and to avoid choosing biased models. Another

panelist responded that many routes of bias can be eliminated by posing the question carefully so as to avoid bias. Bias occurs in data sets, said a third panelist, and analysts should be creative in recognizing bias. For example, sample sets often are haphazard rather than representative and they often exclude representative subgroups. Analysts must be creative about reconstructing, from the available data, the distribution of the true parameter that is really relevant to the risk assessment. Effectively, analysts must model the processes that produce the bias in order to debias.

### *Deterministic Risk Analysis as a Screening Tool*

Panelists supported the idea of using conservative deterministic risk assessment as a screening tool for probabilistic analysis. If the results of that assessment are clearly below the level of concern, then there is no need to do a full-blown probabilistic analysis.

One panelist mentioned that she always performs a deterministic risk assessment first as a screen. When the result is below $10^{-6}$, a probabilistic risk assessment is not performed. When the result is above that threshold, she does a full-blown Monte Carlo analysis. Another panelist said that using point estimates for screening has enabled many facilities in California to avoid costly and unnecessary probabilistic analyses.

### *Model Validation and Data Quality Review*

Panelists briefly mentioned that steps in a probabilistic analysis should also include validation or verification of the model for the sample values, and evaluation of the quality of the available data.

- Whether something affects the tails and does not really affect the central part of the distribution for the result. When this happens, the analysis can produce a very adverse outcome because of bad input assumptions and/or a nonlinear model that a traditional local sensitivity analysis cannot reveal.

- Input variables that do not show up as being statistically significant (for example, in correlations of covariants of an output with an input), but that shift the whole result up or down because they are skewed or because of a nonlinearity in a model. In other words, a variable can appear not to be "sensitive," but can significantly skew the analysis.

The panelist emphasized that analysts cannot rely on local sensitivity analysis and should be very careful in using the preliminary probabilistic assessment to look not only for how the input affects the variance of the output, but also for how it affects the mean and the tails.

Three panelists questioned the value of these proposed approaches. One pointed out that, to do any of these a priori sensitivity analyses, analysts need to know something about the mean and the standard deviation of the inputs anyway, which means they need to know something about the distribution of those data. In generating a standard deviation, an analyst comes close to having a distribution that can be used in a full-blown analysis anyway. And, most of the Monte Carlo analysis software provides a sensitivity analysis at the back end.

A second dissenting panelist pointed out that most analyses are dealing with simple exposure models that have around five or six parameters. For many of these parameters (e.g., body weight), there is no argument about the shape of the distribution. He did not see how sensitivity analysis would play a role in most exposure assessments done in Superfund or indoor air situations.

Another complication with doing sensitivity analyses, pointed out by the third dissenting panelist, is that there can be synergistic effects between variables, particularly when there is more than one pathway where a particular variable may be unimportant as long as another variable goes with it. For example, the amount of time one spends in the yard may be unimportant if the soil adhesion factor is low. The panelist suggested that using numerical experiments to simultaneously vary the variables will be far more informative than simply doing the strawman type of sensitivity analysis one parameter at a time.

Three panelists stressed the value of sensitivity analyses as a tool to focus data gathering. By identifying key variables, sensitivity analysis can provide managers with an important basis for deciding where to focus research dollars to improve distributions.

### 2.1.5 Characterizing the Uncertainty Associated With Use of Surrogate Data

The workshop chair then asked panelists to focus on the following issues statement:

***For some parameters of the exposure equation, site-specific measurements may not be available to determine the probability distributions. In these cases, distributions derived from surrogate data (e.g., national data on body weights) may be used. How do you characterize the uncertainty that has been introduced into the analysis when using surrogate data that are not collected from the population being studied?***

One panelist responded by saying that this type of characterization cannot be done with point estimates or first order random variables. It can only be done by using second-order random variables.

Another panelist pointed out that site-specific measurements are rarely as certain or specific as one would like. Thus, this issue must be approached on a case-by-case basis in terms of deciding whether national data are appropriate for a local situation. National data should not be excluded, as long as the analyst can objectively justify why they relate to the site in question.

One panelist described how he applied national data on fish consumption for a coastal community to a specific site. There were four national candidate data bases. None were ideal, so he ran all four to look at the range of plausible values for the distribution and test whether any of the four data sets made a critical difference to the distribution. He added the weaknesses of all four sets and postulated that the maximum weakness would fall somewhere in that range. Using this argument, he justified that the four data sets provided a reasonable measure of the range for the fish consumption rate at the site.

Building on this idea, another panelist suggested that if any stakeholders had alternative fish consumption data, the analyst could also have run that data set to see whether it made a difference. If it did not, that would be informative. If it did, the stakeholder would have a good point that might influence how the fish consumption rate was ultimately handled in the analysis. One panelist expressed concern that doing this could encourage outlandish proposals from critics who wanted to delay a project.

A panelist pointed out that national data can also help catch an error or identify something inappropriate in local data. For example, if one developed a range of concentrations for a contaminant in the local area that was way outside the background range for a reference population, that might cast suspicion on the local data.

When data from other studies are not that pertinent, said another panelist, then one can characterize the uncertainty in terms of the parameters of that distribution as they may apply to the case in question. Also, there are a set of very effective tools based on Bayesian methods that allow the analyst to have a prior belief (e.g., based on literature values, national estimates) about the distribution of the parameters for the application. This belief may have quite a bit of uncertainty associated with it. These tools include formal techniques that allow an analyst to start off with a prior distribution and then use local site-specific data to come up with an updated or posterior uncertainty distribution. They can be used for the value of information calculations mentioned earlier.

At this point, a panelist introduced the idea of ground-truthing. He pointed out that probabilistic analyses often produce intermediate and final predictions that can be ground-truthed against real data. For the benzene Maximum Achievable Control Technology example, for instance, EPA is predicting the distribution of benzene concentrations in nearby census tracts. There probably are some ambient benzene data, collected at monitors, for refineries that are major contributors to ambient benzene concentrations. Even though these sources may not be the only contributors, it would be very instructive to compare the predicted distribution of ambient benzene concentrations to the observed values. If the predicted distribution was close to ambient values, then EPA could feel some confidence in it. Similar approaches could be used for other values and media. For example, in a body of water, one could take the number people

fishing in the area, multiply by the fish intake, and see whether the result was consistent with the catch. The panelist proposed that using observed data for ground-truthing purposes be an important priority throughout the exposure/risk assessment process.

Another panelist provided an additional example of ground-truthing. He said that data are available from several sources on the population distribution of serum concentrations of fish-related toxicants, such as polychlorinated biphenyls and methyl mercury. An ordinary lognormal fitting provides pretty reasonable fits with about the same geometric standard deviation. This provides confidence that this data set includes the variation and effective intake of those fish liver toxicants, as well as some additional pharmacokinetic variability.

Another panelist pointed out that some of these issues are similar for point estimates. If, for example, one needed a mean estimate of fish consumption and had conflicting studies, it would be exactly the same set of questions: How does an analyst justify supporting one study versus another? How should conflicting data sets be handled? The panelist suggested evaluating the different data sets by applying some weight to each data set. However, if one has a number of data sets and no particular idea which are better, they could be assigned the same probability, but that fact will get lost if one simply crunches numbers. If the data are inconsistent and point in different directions, both the scientist and the decision-maker should know that they are inconsistent and how they affect the bottom line.

### 2.1.6 Collection of Site-Specific Empirical Data for Probabilistic Analysis

Panelists were then asked to focus on the following issues statement:

***If surrogate data are inappropriate for evaluating exposure to the population of interest, site-specific empirical measurements may be necessary. What guidance can be given on the collection of site-specific empirical data to replace the surrogate data used to develop the distribution for a particular exposure parameter? How can you handle subpopulations when developing these data? How can you characterize the reduction in uncertainty associated with the collection of new data?***

The first comment on this issue was that there are several different types of sampling. Stratified random sampling generally is important on a site-specific basis (i.e., sampling randomly among the different strata of the people or organisms that may use the site). The commenter felt that this approach provides the strongest intellectual basis for developing a distribution.

Another approach would be some kind of two-stage sampling, possibly with oversampling in high concentration areas to maximize the information for the areas that might have the most impact.

One panelist stressed the importance of identifying the users for the sampling data and involving them in the sampling design prior to sampling, since different users have different sampling needs. There are primarily two camps of users, the panelist said: those measuring the extent of contamination, and those looking at receptor behavior. From a risk assessment viewpoint, sampling plans should be designed to help elucidate who or what the receptors are, by what pathways they are exposed, and how they are most likely to contact contaminated media. Sampling plans should not simply emphasize extent and distribution.

Another panelist described a situation where engineers wanted to design the sampling plan around delineation to support cleanup decisions, while the risk assessors wanted to design the sampling plan to avoid bias and give the most reasonable risk assessment. In the end, a combination sampling approach was implemented that involved 1) judgmental sampling aimed at finding the edges of contamination, and 2) random, evenly spaced samples more suitable for a risk assessment.

Data quality objectives are used at EPA for various purposes and, suggested one panelist, they perhaps could be generalized to become part of the decision-making connected with exposure assessment and risk assessment. However, the data quality objectives may need multicriteria because there may be more than one objective.

The workshop chair asked if anyone performs a formal analysis to evaluate the behavior of the distributions being formulated based on the sample size and the information collected—i.e., a sensitivity of distribution shape and statistics as it relates to intensity of sample

information gathered. One panelist responded that this would be easier to do when there are phases to the sampling. For example, at one site where three phases of sampling were conducted, the panelist looked at variance after the information from the first phase had been collected and found that more samples were needed. This approach enabled him to design around what the initial sampling revealed.

The workshop chair asked if anything special was being done regarding multistage sampling for probabilistic risk assessment. A panelist responded that one thing being done after multiple sampling runs is to quantify the degree of the lie represented by ordinary standard-error-type uncertainty analysis. He explained that ordinary analyses of uncertainty within a particular data set rely on the implicit assumption that the major source of uncertainty is essentially sampling error (i.e., random fluctuations represented within the data set). But when older measurements are compared to more recent measurements of the same parameters made with newer, more accurate techniques (as can be done in physics, for example), scientists find that the new values wander outside the stated confidence limits for the older values much more frequently than would be expected by chance if the confidence limits calculated under a normal Gaussian distribution were right. This leads to the conclusion that these types of confidence limits are wrong, probably in part because there is unsuspected systematic error in those measurements as well as the random error described by ordinary statistical procedures. Some work is being done to quantify the degree of error that is implicit in various kinds of data. This work is valuable for helping analysts correct these kinds of biases in uncertainty estimates.

### 2.1.7 Estimating Distributions When Empirical Measurements Are Inadequate

The workshop chair asked participants to respond to the following issues statement:

***In some cases, empirical measurements (site-specific or otherwise) for a particular exposure parameter may not be available or may be inadequate to determine a probability distribution. In these situations, should a distribution be estimated to complete Monte Carlo analysis? If so, how? For example, it has been proposed that distributions for these parameters may be***

***estimated via expert judgment or Delphi techniques. If these techniques are used, what factors should be considered in the weight of evidence?***

Panelist response to this issue varied widely. One commenter pointed out that the workshop panelists likely had significantly different views about the nature of uncertainty, which probably would affect their response to this issue. His view was that uncertainty is the state of the lack of knowledge, and that the purpose of the input in a Monte Carlo distribution is to describe the state of the knowledge (i.e., the function/behavior of the inputs within the given model) and not to go much beyond that. At most, a very close extrapolation could perhaps be justified in some cases from the observed knowledge. To the extent that the input is complete, then the results of a Monte Carlo analysis approach a true description of risk. In this sense, the distribution becomes deterministic (i.e., the shape of the distribution is fixed once the input knowledge is complete).

Input information should not be predicted when the input knowledge is incomplete, the panelist continued. Thus, for example, where there are three data points and no information to distinguish among the likelihood of those three points, then the distribution is uniform. If one is more likely than the other two, then one has some sort of a triangular distribution. Analysts who view uncertainty as a construct rather than a state may argue that one can do a lot with a construct. The panelist cautioned against using "pretend data" to do a probabilistic analysis because the "pretend data" will appear real and convey a false sense of confidence. Another panelist supported this viewpoint, suggesting that if distributions have to be invented to do a Monte Carlo analysis, then a point estimate should be used instead.

Another panelist had a different opinion. He stressed that judgment is always present in an analysis. For example, there is judgment about which data to use and whether the particular data are relevant to the particular quantity in question. The relevant question is "How explicit should we be about that judgment in our analyses?" Being honest about that judgment and careful in how the judgment is expressed will likely lead to greater credibility.

The panelist also cautioned that Delphi techniques should not be used as a synonym for the elicitation of expert judgment. There are many techniques for eliciting expert judgment, he

explained. Delphi techniques are a particular method characterized by having a panel of experts who share their opinions in the hope of refining them. Empirical research has shown that all elicitation techniques are likely to lead to overconfidence. However, the Delphi method is particularly likely to lead to overconfidence because of a "group-think" phenomenon that occurs. Thus, other methods are superior because they reduce the degree of overconfidence.

Another problem with Delphi techniques, according to another commenter, is that they can easily be used to get a team of experts to answer a question that does not make sense. This can be avoided by making the experts say why they are providing a particular answer and what the basis is.

Another opinion was judgment must be exercised about when to use a Bayesian approach to think about a problem. This approach can be appropriate when there is an adequate basis for making judgments, but may not be a reliable or practical decision-making approach when the basis is inadequate.

A panelist pointed out that it is important to distinguish between variability and uncertainty. Because variability reflects a natural process, gaps can be filled in by examining available data that relate to the process in question. However, for uncertainty, a whole different class of distributions comes to the fore, including triangular, uniform, and trapezoidal distributions.

One panelist encouraged analysts to think creatively about how to use analogous available data to fill gaps. For example, if one lacks data on a particular species, one could try to find data on a related species. An appropriate reference class can provide an idea of the probabilities. While there clearly is some degree of inaccuracy associated with the use of analogous data, this approach provides an important way of making progress when measurements cannot be done.

A couple of panelists pointed out that, since a point estimate cannot be developed without thinking about range, the issue of using expert judgment is not unique to the probabilistic approach.

One panelist suggested it would be remiss not to mention the technique of maximum entropy as a more objective alternative than expert judgment for developing distributions. Expert judgment and maximum entropy can also be combined, he said. For example, one could select a lognormal distribution based on expert judgment and then choose the lognormal distribution that is maximally entropic.

Panelists mentioned three additional alternatives to using expert judgment: 1) use of default policy values; 2) doing many analyses (a potentially resource-intensive option); and 3) doing no analysis.

Panelists generally agreed that whatever approach was used, it should be transparent. If, for example, judgment forms a large part of an assessment, then that fact should be made clear to decision-makers and the public.

One panelist mentioned that, in her experience, the public tends to think of professional judgment as bias. They do not want it to be considered as information. Therefore, to gain credibility, the analyst must be able to fully support the outcome presented to the public and the rationale for including expert judgment.

Another panelist stressed that the use of judgment in assessing uncertainty is not fundamentally unscientific. Even when there is empirical evidence, judgment is needed to assess the uncertainty. Even the uncertainty in fundamental physical concepts (e.g., the speed of light), which is reported as a standard deviation of the measurement, is judgmental.

A panelist reminded the group that getting expert judgment requires an investment of time and resources. Sometimes this investment would be better spent on simply getting empirical measurements of the quantity in question.

One panelist mentioned that distributions can be portrayed using a scenario-based point estimate approach. In this approach, which has been used in California, the risk is calculated for a variety of plausible scenarios relevant to the exposed population (for example, a homeless

person who does not wash and lives outside, a couch potato, and so on). One important advantage of this approach is that it helps make the results understandable to the public.

### 2.1.8 Characterizing the Effect of a Judgment-Based Distribution on the Tails

The workshop chair asked the group to comment on the following issues statement:

***Can the effect (of using a distribution derived via expert judgment) on the tails of the output distribution be characterized? If so, how?***

One panelist suggested this effect could be characterized by using a point estimate and then using the distribution elicited via expert judgment to see if there was a difference.

Another panelist cautioned that the sort of combinatorial toggling mentioned earlier (trying it with and without) is very difficult because there are infinite families of distributions to choose from and it is hard to do all possible combinations. There are techniques that can incorporate all distributions at once, he said, including probability bounds and a whole tradition of robust Bayesian statistical analyses. These techniques can be used to assess the affect of distribution shape on the tails.

Another panelist mentioned a situation where he used the judgment of three to four experts for a couple of parameters. The experts were asked to explain their judgment. He then ran the model separately for the different expert inputs of the same variables. Three of the four inputs gave roughly the same tail and three gave roughly the same central tendency. Using this method, the source and rationale for the disagreement were transparent.

### 2.1.9 Correlations Among Parameters

The workshop chair asked for comments on the following issues statement:

***Some of the parameters in the exposure calculation may be correlated with each other. Which parameters do we presently know are correlated? Do we know the magnitude of the correlations that exist? These correlations may vary in strength, and the absolute value of the correlations are often unquantified/unquantifiable. If these correlations exist and are moderate to strong, they may have effects on the tails of the output distributions. How should these correlations be accounted for in the Monte Carlo analysis? For example, it has been proposed that one may perform one Monte Carlo simulation with the correlations set to zero and another with the correlations set to some plausibly high value. In this way, the analyst may evaluate the importance of unquantified correlations in the analysis.***

A few panelists stressed that, where possible, dependence should be handled by building it algebraically into the model, so that the parameters cannot vary independently when there is a known dependence. To do this, however, requires knowing what the dependency is. Where possible, the causal mechanism (the reason for the correlation) should be built into the model. In addition, the value of the parameter quantifying the dependence has uncertainty and can also be built into the model.

One panelist pointed out that some dependencies are complex, especially in ecology. For example, sometimes when one looks at both sexes together, the correlation is positive, but when each sex is evaluated separately, the correlation is negative. In such situations, said another panelist, there are probabilistic techniques that can be used to provide bounds on the true dependency structure, but these techniques are not very precise.

There is an idea, said another panelist, that when one does not know the correlation, one should set it to a high value and compute it, and then set it to zero and compute it to see what the range is. This will not work, the panelist said. When multiplying or adding things together, then the extreme correlation is the positive one. But when dividing or subtracting anything, then

the smallest possible correlation produces the widest distribution. This needs to be taken into account.

Panelists mentioned a couple of actual situations where a correlation had a significant impact on the results. One example was an analysis to compute the extinction risk for the spotted owl. When the correlation between mortality of the juveniles and adults is not included, the analysis suggests that the owl is not endangered. When the correlation is included, the analysis indicates the owl is endangered.

One panelist said he used numerical experiments to help resolve whether there might be correlation. However, another panelist cautioned that numerical experiments can only do linear correlations and cannot indicate anything about dependencies more generally.

## 2.2 VARIABILITY/UNCERTAINTY

This topic area opened with a presentation by Christopher Frey on *Quantitative Techniques for Analysis of Variability and Uncertainty in Exposure and Risk Assessment*. Dr. Frey began by defining and contrasting variability and uncertainty. He discussed the different types of uncertainty, issues concerning developing distributions, and the dependencies among variability and uncertainty. He also reviewed approaches to modelling and analyzing uncertainty and variability, and incorporating a discussion of uncertainty and variability into analytical reports. Overheads from Dr. Frey's presentation are provided in Appendix D.

Dr. Frey's presentation was followed by two case studies concerning variability and uncertainty:

- Dr. Timothy Barry reviewed the application of Monte Carlo analysis to an exposure assessment for radon in drinking water. He described the data available for the four exposure model variables that were selected for the analysis. He then explained how the distributions were developed and tested, and how uncertainty was characterized and analyzed. Overheads for Dr. Barry's presentation are included in Appendix D.

- Paul Price presented a case study on applying Monte Carlo analysis to an exposure assessment for a Superfund site. The analysis modelled indirect exposure to TCDD through the consumption of beef from cattle raised down wind of a hazardous waste incinerator. The results were presented as a cumulative distribution of individual doses in an exposed population and the uncertainty in the distribution. The case study illustrated a relatively simple approach to separately evaluate uncertainty and variability in estimates of long-term dose rates. Overheads for this case study are provided in Appendix D.

### 2.2.1 Value of Separating Variability and Uncertainty in Quantitative Analysis (i.e., Second-Order Uncertainty Analysis)

Following the presentations, panelists discussed the value of second-order uncertainty analysis. Several values were suggested:

- Uncertainty analysis provides greater opportunity to be fully objective and to fully explore the quantitative implications of any assumptions than does first order analysis. For example, uncertainty analysis provides information on how imprecise the risk numbers are.

- Making uncertainty explicit enables one to determine whether separate risk estimates agree or disagree.

- Second-order analysis provides more information about the range of the output (e.g., "the best estimate is X with a range of Y to Z"). This provides a range of plausible outcomes for risk managers.

One panelist expressed skepticism about the value of second-order analysis. He argued that because it uses completely subjective descriptions of uncertainty, it appears to add no value compared to other forms of risk assessment.

One panelist said that she had not yet seen any two-dimensional analyses used. She cautioned the group about making recommendations whose wording might suggest that two-dimensional analyses were being recommended as a general principle.

Another panelist suggested that uncertainty should always be addressed, but this did not necessarily have to be done by a formal analysis. For example, uncertainty could also be

addressed by comparing different scenarios and by sensitivity analyses (especially if combined with a qualitative assessment of where variability and uncertainty lie with respect to each of the variables to get an idea about whether one or both of those factors is driving the analysis). Use of interval probability analysis and response surface methods were also mentioned as methods for addressing uncertainty.

A number of panelists supported the idea that uncertainty should be addressed in some fashion. The second dimension keeps the first dimension honest, said one panelist. At a minimum, the analyst should be clear about what is uncertainty and what is variability in the analysis, otherwise both will be inaccurate.

One panelist suggested that performing an interval analysis might provide some insight about the degree to which the uncertainty might contribute to the overall random variation. This approach could potentially serve as a screening tool. If it revealed that uncertainty was not important compared to variability, then a one-dimensional analysis could be sufficient.

Another panelist suggested that a one-dimensional analysis would also be adequate when variability dominated uncertainty by several orders of magnitude, or vice versa; where variability in a population does represent uncertainty for a random individual; and, possibly, when one is looking at uncertainty across an average of a population.

### 2.2.2 Characterizing Model Uncertainty

The workshop chair then asked panelists to address the following issue statement:

***How can one adequately characterize the uncertainty associated with the selected conceptual and mathematical models? Can all types of variability and uncertainty be analyzed using techniques such as Monte Carlo analysis?***

Panelists had a variety of responses to this question. One opinion was that uncertainty was more difficult to characterize for a conceptual model, such as a dose-exposure model, and easier to characterize for a mathematical model, such as describing body weight distribution.

One panelist suggested a posterior analysis approach to characterizing uncertainty, which he felt was less subjective than the two-dimensional approach. This approach works as follows. Analysts can readily determine which inputs have a larger or smaller effect on the model by evaluating the sensitivity of the individual inputs. And, because analysts have a qualitative sense about which inputs to the variables are driven by uncertainty and which are driven by variability, they can then have a good sense of whether the inputs with the largest effect are driven by uncertainty or variability or both. This provides a qualitative way of identifying the degree to which uncertainty affects the final outcome. This approach could be taken one step further by quantifying uncertainty, albeit highly subjectively. Conceptually, this approach is not different from quantifying uncertainty at the front end, but it does not pretend to give a numerically precise estimate of the variability on the tail end. The panelist preferred this method to others that pretend to be more objective.

Another panelist seconded this idea, especially in situations where there is little data. She suggested that the model could also be run with a few different point estimates for the uncertain parameters.

Another panelist said that the broadest mathematical framework currently available to address this type of problem is to use all variables as second-order random variables. He felt that this is a powerful and important mathematical method, but that there is resistance to its use because it is new. He expected use to increase over time as people become familiar with it.

A number of panelists expressed concern that uncertainty analysis could mislead decision-makers about the magnitude of the uncertainty because there are more uncertainties than can be captured by an uncertainty analysis. This sentiment was expressed by one panelist who said that assessors are being encouraged to use Monte Carlo analysis, which is being "sold" as a software package that will readily provide distributions from which one can pick off a predetermined regulatory decision point (such as the 95 percentile) and compare it to the point estimate. He

was concerned that the technique had been oversimplified and that many people had no idea about the uncertainty involved in using it.

Agreeing that it is not possible to analyze all uncertainty, another panelist suggested that the analysis focus on the important uncertainties and clearly disclose what set of uncertainties the model attempts to represent and what it does not. He suggested that analysts be vigilant in identifying the full range of types of uncertainty and variability impinging on an analysis.

The panelist defined uncertainty as "the distribution of the likelihood that the analyst is wrong by various amounts," including model error, measurement error, and use of nonrepresentative data. For example, when samples are collected haphazardly, rather than by utilizing a stratified random design, the resultant estimate is more uncertain and could be biased. Also, uncertainty can be introduced when the population to be modelled differs from the population from which data were obtained. Further, the standard error calculation does not take into account uncertainties resulting from calibration of analytical equipment. Finally, several kinds of model errors can affect the accuracy of a uncertainty analysis.

Another panelist mentioned that systematic error is one of the key forces in uncertainty. Statistics can be used to estimate random variability, but they cannot be used to make inferences about systematic error. There is a danger, therefore, that statistics could be used to invent data when it is missing.

In response to a question about how Monte Carlo analysis might account for different conceptual representations of natural phenomena (e.g., fish bioaccumulation of aquatic contaminants), a panelist suggested that model discrimination could be used to compare the predictive power of the alternative models. Another approach would be to include each alternative model in the uncertainty analysis. Alternatively, one could compare models to data and, if there was a good fit, give greater weight to the samples that would go with that model in the overall uncertainty analysis.

One commenter expressed concern about the intentional or unintentional human bias that may be introduced into an assessment because of the particular stake that the risk assessor

or the parties he or she represents have in the outcome. Risk assessments closer to the default or point estimate approaches are easier to compare to see what assumptions were made. The commenter predicted that, as analyses start to include second-order considerations, views would diverge and could not be compared because subjectively based second-order estimates of uncertainty are all equally valid since they each are an accurate reflection of the analyst's judgement. He expressed concern that analyses based on opinions may result in manipulation even when intentions are good. In other words, they offer a real potential for bias that cannot be corrected.

Another panelist responded by suggesting that those flaws would hopefully be uncovered as the individual risk analyses were subjected to review by a variety of viewpoints. The challenge for risk analysts is how to design analyses that make the assumptions clear so that they can be inspected, and that make the uncertainties explicit so they can be discussed and reviewed. Panelists discussed the appropriateness of using interval analysis, as opposed to a distributional approach, to represent uncertainty. One panelist argued strongly that interval analysis was appropriate when the uncertainty was large and the variability small, whereas Monte Carlo analysis was appropriate when the uncertainty was small and the variability large. In his opinion, risk analysis is a subdiscipline of probability theory and is focussed on frequencies, because the real issue is about how many people are going to be affected by exposures. This is a frequency issue rather than a subjective probability issue.

He explained that frequency distributions cannot be known precisely. That being the case, expressing these distributions as probability distributions simply yields a distribution of distributions. A better approach is to have a range of distributions, which can be easily displayed as an interval probability. An interval probability effectively means the true frequency or probability distribution is somewhere in that range, but the analyst is not sure where because the uncertainty is large. If one read off the 75th percentile, for example, one would get an interval that represents the uncertainty about that percentile. When uncertainty gets small, however, the range contracts and becomes a single distribution so use of probabilistic analysis becomes more appropriate.

Strictly speaking, he continued, large uncertainty is not a probability and therefore should not be combined analytically with variability, which is a probability. If a distributional approach is used when uncertainty is large, then the analyst's subjective feelings about the problem can enter into the analysis. The commenter suggested that, as a general rule, uncertainty and variability should be treated differently from a computational standpoint, otherwise the tails become devalued in an obvious way.

A number of panelists held contrasting views. One panelist pointed out that all analysis is subjective; different approaches simply differ in the degree of subjectivity. Another panelist said that many people do think that uncertainty can be represented with probabilities. He said there was substantial evidence to suggest that Monte Carlo was a good approach to doing this. He felt that the real issue is whether probability can be used to represent judgment. A third panelist commented that interval analysis is not very informative, therefore judgment about the bounds is needed. He suggested that interval analysis be one tool that could potentially be utilized in a tiered approach.

Panelists briefly discussed methods for quantitating a distribution of uncertainty versus variability. One innovation is to use an exponential distribution form (rather than a Gaussian form), which spreads the tails out to a greater extent than the midpoints. The likely unsuspected systematic error is another distribution that can be readily put into an analysis. However, it does not capture model error. The parameter for the exponential distribution is calculated by making an analogy between the kinds of measurements one has and the kinds of measurements whose expedience has been summarized by the exponential formula.

Panelists then discussed whether variability and uncertainty should be tracked and evaluated separately during the analysis. This must be decided individually for each separate variable, said one panelist. The only general principle is that one *can* separate variability and uncertainty. Another panelist suggested the following principle: It is often useful to separate variability and uncertainty when that provides more accountability and transparency about the assumptions. Another suggestion was the decision about whether to separate uncertainty and variability would depend on how the question was defined. In general, panelists agreed that this

decision depends on context, that it is often useful to track these separately, and that a conscious decision about whether to separate them should be made after the problem is formulated.

## 2.3 PRESENTING RESULTS

Discussion concerning the presentation of results from a Monte Carlo analysis was guided by two presentations:

- Thomas McKone reviewed a number of issues concerning presentation: What should be presented; how variability and uncertainty can be characterized in a presentation; comparison of Monte Carlo results to point estimates; how to characterize the results of a sensitivity analysis and the stability of the tails; how to present the results of an expert elicitation; and incompatibility. Overheads for Dr. McKone's presentation can be found in Appendix D.

- Max Henrion's presentation focussed on communicating and documenting uncertainty in risk analysis. He illustrated the use of a software model that utilizes hierarchical influence diagrams and integrated model documentation to document and communicate uncertainty. He also emphasized the importance of sensitivity analysis to identify the relative importance of sources of uncertainty. Finally, Dr. Henrion summarized eight reasons to model uncertainty. Overheads for Dr. Henrion's presentation are included in Appendix D.

Panelists then discussed issues concerning presenting the results of Monte Carlo analyses.

### 2.3.1 Identifying and Understanding Audiences for the Presentations

Panelists agreed that presentation of results needs to be tailored to the audience. Two primary audiences for the results of probabilistic risk assessment are the public and decision-makers. Ideally, analysts should take time prior to presenting the results (preferably before even starting the analysis) to get to know the audience's needs and expectations regarding the results, as well as their overall knowledge about risk assessment in general and the Monte Carlo process in particular. This information will help analysts tailor their presentation to the audience. One panelist proposed the idea of having a communications specialist involved in all phases of the

assessment, beginning with problem formulation, to improve communication between the scientists and the public and to better target the presentation of results.

### 2.3.2 Building Trust and Understanding

Panelists shared their recollections about a risk communication study[2] that examined how the public reacted to various hypothetical risk estimates presented by EPA. One experiment found that people reading a simulated news story containing a range of EPA risk estimates agreed that the Agency's discussion of how much the risk might vary made it seem more honest (66%) and disagreed that this discussion made the Agency seem less competent (59%). Preliminary regression analysis of other experimental data found that people reading such stories, compared to people reading stories with point estimates of risk, rated EPA as less competent at risk assessment and environmental management, but rated the story's truthfulness higher. People in these studies had at least some college education; the researchers had no data to suggest whether people with less education would have different reactions to being presented with a range of risk estimates.

A panelist pointed out that process drives perception to a large extent. Stakeholders[3] will be more trusting when they are involved before and during the analysis (to frame the problem, etc.) to provide input into the problem structure. Also involvement provides a chance for the analysts to understand the audience's needs vis-a-vis presentation and to educate stakeholders and decision-makers as the process evolves so that they can have a better understanding of the results. A number of panelists supported the idea of stakeholder involvement where possible, although they acknowledged that it is not always possible.

---

[2]Johnson, B.B., and P. Slovic. 1995. Presenting uncertainty in health risk assessment: Initial studies of its effects of risk perception and trust. Risk Anal. 15(4):485-494.

[3]A "stakeholder" is anyone who has an interest or stake in the outcome of the process in question (in this case risk assessment).

The scenario-form of analysis described earlier (where risks are estimated separately for a variety of exposure scenarios of relevance to stakeholders) was mentioned as one approach whose results particularly lent themselves to stakeholder understanding.

One panelist pointed out that a broader, long-range educational process is occurring that has relevance to the public's understanding of and attitude toward risk information. For example, risk analysis originally focussed on the issue of "is it safe," as reflected in the Delaney clause. But many people now understand that safety cannot be absolute—that it is a question of degree and that risk management is a question of tradeoffs between the costs and benefits. The panelist thought that, as a result of this long-range educational process, many members of the public are now ready to hear about uncertainty. He also speculated that the concept of uncertainty would become increasingly accepted and understood over time as it is applied in various areas of relevance to the public, such as in earthquake prediction.

In one panelist's experience, risk managers and the public generally can understand that exposure is a distribution rather than a single value. However, because of poor risk communication, many risk managers and the public think the results of a risk analysis are an actual estimate of risk at a site and that the number cannot be exceeded. They often are surprised to see how a point estimate or the Reasonable Maximum Exposure (RME) compares to the results of a distributional analysis. Risk managers are pleased to find there is maneuverability.

The discussions suggested a variety of levels of familiarity with Monte Carlo analysis among risk managers. For example, it was one panelist's impression that a number of risk managers are not very familiar with the technique. Therefore, it is important to start educating them about probabilistic analysis so they can better make use of the results.

Another panelist had the opposite impression. In her experience, many risk managers are already familiar with the technique. They want uncertainty expressed in a clear and useful way. They want to know, at least qualitatively, what the limitations are, where the weaknesses are, how certain the analyst is of the estimate, and whether the analyst can support it. She said that risk managers will not make a risk management decision without this type of information.

At least one panelist's experience suggests that many decision-makers want a relatively simple system/tools to distinguish between a good Monte Carlo analysis that furthers decision-maker process and a bad (technically flawed or irrelevant) Monte Carlo analysis. This could be in the form of, for example, a checklist, flow chart, set of evaluations, case study, or a casebook of "exemplary Monte Carlo analysis").

### 2.3.3 Presentation Formats

Panelists discussed the types of formats they have used to present the results of a probabilistic analysis. They agreed that entirely different types of reports should be used for scientific and nonscientitic audiences. Reports for the scientific community are usually very detailed; while descriptive, less detailed summary presentations (e.g., box and whiskers, simple tables) and key statistics with their uncertainty intervals are more appropriate for nonscientists.

One panelist suggested a tiered approach to presenting results, where the level of detail increases with each successive tier. For example, the first tier could be a one-page summary that might include a graph or other numerical presentation as well as a couple of paragraphs outlining what was done. The next tier could be an executive summary, and the third tier could be a report. Decision-makers probably would never read a report, but they might want it presented. A panelist suggested that the Congressional Research Service, Office for Technology Assessment, or other agencies that present info to high-level decision-makers might have some valuable models for presenting information that could be of use in designing presentations formats for risk information.

One panelist pointed out that simplifying and highlighting results for a less-detailed presentation involves making a decision about which numbers are most relevant for decision-making. Analysts need to be sensitive to this value issue. To avoid imposing their own biases, they should confer with risk analysts about which type of results are most important to decision-making. Analysts can ask the managers, for example, "Do you as an risk manager want to be X% confident that the relative risk for the Yth percentile individual is below Z?" This type of precision will help maintain consistency. A lack of precision will lead to inconsistency. For

example, if a risk manager simply says "Keep the risk below $10^{-6}$," the analyst may do this by playing with either the variability or uncertainty dimensions. The panelist suggested that, ideally, the information presented to risk managers should be comprehensive and clear enough to enable them to make as informed a choice about risk management as they would have if they themselves had gone through the analytical process.

A panelist suggested that tables of what is at the 90th percentile and above would be useful for risk managers because they use those sorts of things when considering acceptable risk. She said the questions given in the Bloom et al. (1993) report (see Appendix F) largely reflected the types of questions that risk managers typically asked.

Another panelist suggested that sensitive populations should be highlighted when presenting results. Impact on sensitive populations is something risk managers need to consider. Separating these results out also will facilitate presentation of the results to the public.

One panelist mentioned that presenting risk information often is done in an adversarial context. In that case, the presenter needs pointers back to the sources of the information, preferably showing how it is ground-truthed. Another panelist praised a hypertext system presented during the workshop that allowed the user to obtain various levels of information, including references, about an item simply by clicking on it. She also noted that presenting the sensitivity analysis had proven particularly useful with risk managers in her state.

# SECTION THREE

# PRINCIPLES AND RECOMMENDATIONS

A number of key themes and areas of agreement emerged during the workshop discussions. At the close of the second day, the workshop chair highlighted these areas and tasked workgroups of panelists to develop statements that incorporated and, as appropriate, expanded on the ideas and opinions expressed during the earlier discussions. The workshop chair also developed strawman statements of principles and conclusions.

On the final day of the workshop, the panelists reviewed, discussed, and refined the strawman principles and workgroup statements. The statements were then circulated to panelists for review prior to publication. The final principles statements developed by this process are described in Sections 3.1 through 3.3, and the recommendations are listed in Section 3.4 below. They represent the consensus or near consensus of the workshop panelists. Any key areas of dissension are noted.

In addition, panelists reviewed and discussed the workgroup reports during the final day of the workshop. Workgroup chairs then finalized their reports after the workshop in light of these discussions. The final reports are reproduced in Appendix E. Panelists did not comment on the final reports, therefore, they should not necessarily be construed as consensus.

## 3.1 CROSS-CUTTING PRINCIPLES AND CONCLUSIONS

Interwoven into the discussions of input data/distributions, variability/uncertainty, and presenting results, were a number overarching discussions about when and how to use Monte Carlo analyses—cross-cutting issues that are intimately connected with the workshop's three primary topic areas.

Decisions about when and how to use Monte Carlo analysis include policy considerations, which were outside the scope of the workshop. They also involve practical technical and communication considerations, which were discussed. Exposure assessments can have a range of possible objectives and can be conducted using a variety of approaches ranging from simple to complex. For example, analyses may include "full risk assessments" where the analyst examines the full range of uncertainty as well as narrower "safety assessments." Panelists articulated the following three principles regarding the application of Monte Carlo analysis to exposure assessments.

### 3.1.1 Defining the Objectives of the Assessment

***Exposure assessments utilizing Monte Carlo analyses should begin with a clear question or questions. These should be developed through discussions with the risk manager and should take into account the purpose of the assessment and, where appropriate, the concerns and input of the interested parties.***

Clearly defining the purpose of an assessment is a critical first step, since the appropriate scope depends on the purpose. At a hazardous waste site, for example, the approach to the analysis may be very different depending on whether the goal is to protect a specific population or to achieve a certain cleanup level.

The purpose of an assessment, in turn, depends on how the "problem" the assessment will address is formulated and how the results will be used. For this reason, input from the risk manager who will use the assessment results is critical. Also, input from stakeholders (e.g., the public) at this stage can provide important insight into the problem and helps ensure that the results of the analysis will be responsive to stakeholder concerns and questions.

Workshop discussion relevant to defining the objectives of tiered approach can be found in Section 2.1.3.

### 3.1.2 Tiered Approach to Utilizing Monte Carlo Analysis

***Where possible and appropriate, exposure assessments using Monte Carlo analyses should proceed using a tiered strategy.***

As noted above, the level of sophistication of exposure assessment and associated uncertainty analyses can range from simple to complex. A tiered approach that incorporates Monte Carlo analysis acknowledges this fact by specifically defining a number of approaches, or levels, that begin with a relatively simple approach and progress stepwise to increasingly sophisticated approaches. Because Monte Carlo analysis can be a resource-intensive activity, the level of sophistication should be appropriately tailored to the goals of the analysis (as articulated under the principle described in 3.1.1 above). In some situations the simplest approach (i.e., screening-level analysis) may provide adequate information to fulfil the purpose of the assessment. In other cases, a more sophisticated approach will be needed, and certainly some situations will require use of the most sophisticated approach.

In many situations, after defining the goals of an assessment, it may be obvious which level or tier is most appropriate for the assessment. In other cases, however, this will not be immediately obvious. In this case, a tiered approach offers a tool that analysts can use to systematically discern the most appropriate level of analysis. For example, the first tier typically would include developing point estimates of risk to a high-end individual. If the point estimate of high-end risk is lower than the regulatory level of concern, then the analysis may be complete. Additional tiers could provide, for example, qualitative evaluation of model and scenario sensitivity, quantitative sensitivity analysis of high-end or mid-range point estimates, and, at the highest tier, full quantitative as well as qualitative characterization of uncertainty and the importance of components contributing to the uncertainty. A review of previous applications of Monte Carlo analysis in similar situations may be useful in defining the appropriate level of analysis.

A workgroup chaired by Tom McKone developed a statement describing a potential tiered approach to uncertainty and variability analysis in exposure assessment. This statement is

reproduced in Appendix E. Note that this suggested approach is geared toward regulatory programs. Monte Carlo analyses for other purposes might proceed differently.

Tiered approaches are gaining wide acceptance by states, federal agencies, and industry as reflecting cost-effective strategies for environmental management. Examples include the ASTM (American Society for Testing and Materials) Risk-Based Corrective Action Program, EPA's Soil Screening Methodology,[4] the Massachusetts Contingency Plan, and Nuclear Regulatory Commission Decontamination Plan.

Workshop discussion relevant to use of a tiered approach can be found in Sections 2.1.3 and 2.1.4.

### 3.1.3 Formulating the Conceptual and/or Mathematical Model

***When formulating a conceptual and/or mathematical model for the exposure analysis, identify and consider the various options for formulating the model. Document how and on what basis the model was formulated.***

Model formulation typically involves discussions among the appropriate technical people and, where appropriate, input from other interested parties. Influence diagrams (which illustrate relationships among the components of the analyses) or similar approaches are useful for examining the relative importance of various alternative models or model components. The analyst should document why particular models or model components were selected among the alternatives.

---

[4]U.S. EPA. 1996. Soil Screening Guidance: User's Guide. EPA/540/R-96/018. Office of Solid Waste and Emergency Response, Washington, DC.

U.S. EPA. 1995. Soil Screening Guidance: Technical Background Document. EPA/540/R-95/126, PB96-963502. Office of Solid Waste and Emergency Response, Washington, DC.

In some cases, it may be appropriate to include alternative conceptual or mathematical models in the exposure analysis if there is an insufficient basis for selecting among them. The outputs of these alternative models can be incorporated formally into the overall uncertainty analysis.

## 3.2 DERIVING AND USING INPUT DATA AND DISTRIBUTIONS FOR MONTE CARLO ANALYSIS

### 3.2.1 Determining Whether To Develop Distributions for Some or All Variables

***Specifying distributions for all or most variables in a Monte Carlo analysis can be useful for exploring and characterizing the full range of uncertainty.***

Some workshop panelists have found it useful to include distributions for all or most of the variables in an analysis. When used in conjunction with sensitivity analyses, this approach enables the analyst to explore the possible ranges of uncertainty and the relative importance of the variables to the overall uncertainty. Such information can be useful, for example, to direct future data collection efforts to reduce uncertainty. The decision about whether to include distributions for all variables generally is not affected by computational limits (i.e., current computers and software usually can handle the task).

Panelists expressed a range of opinion about whether it is appropriate to develop distributions for variables for which little data are available. A number of panelists cautioned against "inventing" distributions when the input data are incomplete. They recommended that a point estimate be used instead of a distribution in such situations. Other panelists argued that judgment is always present in exposure analyses, even in point estimates, therefore a distributional approach can have validity even when input knowledge is limited. One suggested approach to resolving this problem was to employ distributions that describe what is known about the existing data, but that do not necessarily attempt to describe the true, but unknown, underlying distribution. All panelists agreed, however, that whether or not a full distributional

approach was used, the analyst should be as clear and explicit as possible about where and how judgment was used in the analysis. (See Section 3.2.5 for more discussion of this issue.)

Panelist discussion on this point can be found in Sections 2.1.4 and 2.1.7.

***Point estimates may be combined with distributions in a Monte Carlo analysis.***

From a computational standpoint, a Monte Carlo analysis can include a mix of point estimates and distributions. Individual decisions to combine point estimates and distributions for a specific Monte Carlo analysis reflect a combination of practical considerations (i.e., the costs associated with obtaining the information needed to derive distributions), philosophical differences regarding how uncertain variables should be included in the analyses (see above principle), and the purpose of the exposure assessment. Numerical experiments and sensitivity analyses can be helpful in evaluating the effects of these combinations on the final result. A decision on whether to include point estimates along with distributions in a Monte Carlo analysis should reflect one of two considerations:

- Sensitivity analysis—at some level—has indicated that including point estimates does not affect the overall analysis. Numerical experiments and/or formal sensitivity analyses can be especially useful for determining which, if any, of the variables could be represented by point estimates without greatly affecting the overall results.

- A scenario-based approach[5] is being employed in which the uncertainties associated with exposures for specific "fixed" scenarios are being examined. This approach is sometimes taken when data are limited or when the analyst is addressing a narrow or focused risk management question. Such approaches may not need to incorporate the full range of uncertainty into the Monte Carlo computational framework.

[5]An exposure scenario is a set of facts, assumptions, and inferences about how exposure takes place that aids the exposure assessor in evaluating, estimating, or quantifying exposures. In a scenario-based approach, certain exposure assumptions or factors may be held constant (e.g., exposure duration, frequency, or concentration). Results obtained from one scenario may be compared to those of another scenario that uses a different point estimate for these variables.

***Numerical experiments are useful tools for examining contributions that individual variables make to the overall uncertainty.***

With the advent of faster desktop personal computers and more efficient simulation techniques, it is possible to evaluate a number of "what if" cases that can generate insight into which assumptions or variables significantly affect the answer. Similarly, numerical experiments (i.e., in which the analyst explores the results obtained with a variety of input values, either individually or in combination) can be used to identify key sources of variability and uncertainty with respect to the assessment endpoint. A workgroup composed of Christopher Frey and Scott Ferson prepared a statement regarding the use of numerical experiments in Monte Carlo analysis. This statement is reproduced in Appendix E.

***When data are unavailable for an important variable in an exposure model, it may be useful to define plausible alternative exposure scenarios to incorporate some information on the impact of that variable in the overall assessment of exposure.***

Scenario-based Monte Carlo analyses may be a useful approach for exposure assessment when the lack of data make development of a distribution very uncertain for a given exposure variable and/or when such approaches would help facilitate communication with risk managers and the public. In such cases, it may be relatively simple to develop scenarios that use different values for a key variable in the model.

For example, in an assessment of the exposure to malathion sprayed in an urban aerial setting in California, the California Department of Health Services (CDHS) developed discrete scenarios that varied the length of time one spent outdoors playing on or touching contaminated surfaces (California Department of Health Services, 1991; Marty et al., 1994). At the time the risk assessment was conducted, there were no available data on the amount of time people spend playing on playground surfaces, playing soccer, and so forth. The CDHS, in consultation with an expert scientific review panel and community members, decided to assess a variety of exposure durations. Exposure duration was directly proportional to both the amount of malathion-containing material one contacted via dermal exposure and inadvertent soil ingestion. By varying

the length of time in contact with contaminated surfaces, the assessment was able to portray the uncertainty and variability in that particular parameter without having to try to develop a distribution in the absence of data.

When scenario-based approaches are used, it should be recognized that these approaches do not provide complete analyses of the uncertainties associated with exposures and that the variables that are "fixed" may be important sources of uncertainty. Therefore, it is prudent to present the results using more than one value for the variable (perhaps two or three across a range of plausible values) to prevent the perception that the variable is not contributing to the variation or uncertainty of the analysis and to provide a more complete picture for risk managers and the public.

The use of scenario-based approaches can facilitate communication with risk managers and the public. Therefore, decisions to use scenario-based approaches are often based on the need to communicate information to these groups. Often the scenarios are tailored to specific concerns raised by the public or by risk managers. In such cases, the scenario-based approach is both a communication and an analytical tool; it may or may not be accompanied by a more complete analysis of uncertainty that could help put the selected scenarios into perspective.

Scenario-based analyses also provide a form of sensitivity analysis. They can help illustrate how the exposure models behave given different assumptions. This is especially useful when it is not possible to validate the models. In these situations, the scenario-based approaches can help engender confidence in the models by showing that they behave in reasonable ways. This is sometimes difficult to do with probabilistic analysis alone.

***The analyst and risk manager should continually review the bases for "fixing" certain parameters as point values to avoid the perception that these are indeed constants that are not subject to adjustment.***

Once certain variables becomes "fixed" for a particular application, they may be viewed as standard default values that can be used across applications. Workshop participants cautioned

against this. For example, variables that are relatively unimportant to the overall uncertainty in one case, may be very important in another. Also, while a scenario-based approach may be desired in one application, a more complete uncertainty analysis may be appropriate for another. Thus, the basis of the decision to include specific point estimates along with distributions should be reexamined for each subsequent application to ensure that the overall approach meets the objectives of the exposure assessment.

### 3.2.2 Utilizing Sensitivity Analyses

***Sensitivity analyses can be helpful in identifying parameters that have the most influence on the final result, and can aid in focusing data-gathering efforts. This is especially useful when it is costly to obtain the data for a particular distribution.***

Sensitivity analyses can be helpful for identifying the relative importance of the variables in terms of how they influence the outcome. This can aid in focusing data-gathering efforts. A key motivation for conducting a sensitivity analysis is to identify those model assumptions (both parameter values and conceptual/structural formulations) that most affect the model results and thus the decisions derived from them. Once identified, these more "influential" assumptions and parameter values are prime candidates for further studies (e.g., model analysis, research, experimental studies and field data collection programs) to reduce uncertainty and improve the basis for decision. Three broad classes of sensitivity analysis methods can be identified ranging from simple to more complex:

- Methods that compute the direct response of the model to changes in input values or assumptions—these methods generally involve simple perturbations of the model, and are usually employed prior to more sophisticated evaluations of uncertainty.
- Methods conducted as part of the uncertainty analysis, often through further analysis of the simulation results.
- Decision-driven methods that assess the impact of uncertainty in input assumptions on pending decision and the potential loss (i.e. costs and benefits) associated with them.

A workgroup chaired by Mitchell Small identified a hierarchy of methods for sensitivity analysis based on these three broad classes (see Appendix E). For workshop discussion that led to the formulation of this principle, see Section 2.1.4.

### 3.2.3 Using Surrogate Data

***Surrogate data can be used to develop distributions when the surrogate data can be appropriately justified. The analyst should identify (where possible) and evaluate the factors that introduce uncertainty into the analysis. In particular, attention should be given to biases that may exist in the data sets.***

Data used for distributions include physical, chemical, and biological phenomena that affect the fate of chemicals, activities and characteristics of receptors, and forms or shapes of distributions. Ideally, these data are obtained directly by studying the particular site or situation in question. Often, however, an analyst does not have adequate situation-specific data. In such cases, additional data from comparable situations may be used as a surrogate. Typically, surrogate data come from other studies.

A number of techniques can be used to judge the applicability or usefulness of surrogate data and/or to adjust the data. Examples include comparisons among the pool of potential surrogate data sets, adjusting the data to conform to known differences in processes or biases, and use of Bayesian methods (see Section 3.3.3). Resampling methods may be useful for identifying the uncertainties associated with applying data sets developed for large populations to small populations.

Discussion regarding this issue can be found in Section 2.1.5.

***Whenever possible, develop data—even limited data—to help ground truth the distribution based on surrogate data.***

The use of surrogate data to develop distributions can be made more defensible when case-specific data are obtained to check the reasonableness of the distribution. Collection of such data can provide an important reality check on the analysis. In Monte Carlo exposure assessments involving the use of surrogate data, a balance should be sought between using readily available surrogate data sets and case-specific measurements and observations.

***When alternative surrogate data sets are available, care must be taken when selecting and/or combining sets.***

When multiple surrogate data sets are available, analysts must decide whether to use only the single best data set or to combine the data sets. When the alternative data sets are all reasonable (i.e., one is not clearly the best), the analyst should consider the appropriateness of combining across data sets to develop one distribution. Alternatively, a scenario-based approach could be used where the analysis is run separately for each data set. Workshop participants generally agreed that further guidance is needed about combining data sets, including how to handle conflicting data sets.

### 3.2.4 Obtaining Empirical or Site-Specific Data

***When obtaining data for developing input distributions, particular attention should be given to the quality of information at the tails.***

Management decision are often made utilizing information at the tails of Monte Carlo output distributions (e.g., 90th or 95th percentile) as well as central values. However, the quality of information at the tails of input distributions may not be as good as the central values. The analyst should pay particular attention to this issue when developing input distributions to help ensure the reliability of the estimates at higher percentiles of the output distributions.

*When developing empirical data for distributions, the basic tenets of sampling for exposure assessments should be followed.*

As a general rule, the development of data for use in distributions should be carried out using the basic principles employed for exposure assessments. These include:

- Receptor-based sampling in which data are obtained on the receptor(s) or on the exposure fields relative to the receptor(s). Typically, such sampling takes into account where, when, and how the receptors would be exposed, either by virtue of receptor activity or exposure concentration spatial/temporal profiles.
- Sampling at appropriate spatial or temporal scales using an appropriate stratified random sampling methodology.
- Using two-stage sampling to determine and evaluate the degree of error, statistical power, and subsequent sampling needs.
- Establishing Data Quality Objectives for the exposure assessment.

These subjects have been addressed in numerous EPA and other publications on exposure assessment. In addition, a workgroup chaired by Teresa Bowers prepared a statement on common sampling-related issues that arise when conducting exposure assessments, including Monte Carlo analyses, concerning soils. This statement is provided in Appendix E.

### 3.2.5 Utilizing Expert Judgment Within a Monte Carlo Analysis

***Depending on the objectives of the assessment, expert judgment can be included either within the computational analysis by developing distributions using various methods or by using judgement to select and separately analyze alternate but plausible scenarios. When expert judgement is employed, the analyst should be very explicit about its use.***

The elicitation and use of expert judgement in decision-making is not novel for Monte Carlo analysis. The use of expert judgment has a long history and methods for eliciting expert judgement are well-recognized. Expert judgement is typically used to some extent throughout all exposure assessments.

Panelists agreed that expert judgment ***can be*** included within a Monte Carlo analysis from a computational standpoint. However, they disagreed about when and how distributions based on expert judgement ***should be*** included in a Monte Carlo analyses. Key issues of disagreement concerned:

- The extent to which expert judgment is used to derive distributions.
- The manner in which such judgments are made.

Panelists' differences reflected differing philosophies on what kinds of information should be included within the computational framework of a Monte Carlo analysis. The case for using expert judgment to derive distributions is that these distributions reflect bounds on the state of knowledge, provide insight into the overall uncertainty, and can be used to evaluate the importance of the individual variables.

The case for keeping expert judgement outside the computational framework is primarily based on a desire to qualify or distinguish information elicited by experts from other types of information used in the Monte Carlo analysis. In particular, some workshop participants expressed the following concerns:

- Distributions based exclusively or primarily on expert judgement reflect the "opinion" of individuals or groups, and such opinions can vary and be subject to biases of the individual(s) constructing the distributions.
- Distributions based on expert judgement may be viewed as equivalent to those based on hard data, thus giving them greater credibility within the analysis than may be warranted.

Panelists did agree that however expert judgment was used, the analyst should be explicit about its role in the analysis.

Max Henrion and Clark Carrington prepared a statement about the role of expert judgment in exposure assessment. This is reproduced in Appendix E. Discussion on this issue can be found in Section 2.1.7.

### 3.2.6 Dealing With Correlations Within a Monte Carlo Analysis

***Correlations among input variables can be important and can be handled computationally within a Monte Carlo analysis.***

Several different types of correlations may be encountered during a Monte Carlo analysis. Some (e.g., body weight and body surface area) are obvious. Others (e.g., various dependencies among physical processes or receptor-related activities) may be less apparent. Panelists agreed that correlations can have a significant influence on the outcome of a Monte Carlo analysis, therefore, care must be given up front to identifying correlations that may be important. Scott Ferson and Christopher Frey prepared a statement, reproduced in Appendix E, on methods dealing with correlations. Discussion on this issue can be found in Section 2.1.9.

## 3.3 EVALUATING VARIABILITY AND UNCERTAINTY

### 3.3.1 Defining Variability and Uncertainty

***The concepts of variability and uncertainty are distinct. They can be tracked and evaluated separately during an analysis, or they can be combined within the same computational framework. Separating variability and uncertainty can be useful to provide greater accountability and transparency. The decision about whether to track them separately should be made on a case-by-case basis.***

*Variability* represents the heterogeneity or diversity in a well-characterized population. Variability is a bounded characteristic of the population that may not be reducible through further measurement or study. Variability is sometimes referred to as "Type A Uncertainty." Variability may have some uncertainty associated with it. For example, if only a subset of the population is measured or if the population is otherwise undersampled, the resulting measure of variability may differ from the true population variability.

*Uncertainty* represents 1) lack of information or knowledge about a phenomenon, or 2) lack of the ability of a model to represent the process of interest. Uncertainty is sometimes referred to as "Type B Uncertainty." It is sometimes reducible through further measurement or study.

A variable may reflect primarily variability, primarily uncertainty, or both variability and uncertainty. It is useful for the analyst to recognize and distinguish between the variability and uncertainty features of the variables. This can be helpful, for example, when deciding where new information should be obtained to reduce the uncertainty in the results.

Variability and uncertainty *can* be separated computationally during the analysis. Separation can provide greater transparency and accountability. The extent to which variability and uncertainty *should* be separated computationally during an analysis will depend, in part, on the needs of the project. The analyst should make a conscious decision about whether to separate them after the problem is formulated.

Section 3.1.2 describes a tiered approach to conducting a Monte Carlo analysis within the framework of an exposure assessment. The suggested tiered strategy is useful for considering how to proceed from simple to more sophisticated analyses involving variability and uncertainty. A workgroup chaired by Tom McKone prepared a report describing a suggested tier strategy. This is included in Appendix E.

Panelists' comments on the issue of separating variability and uncertainty can be found in Sections 2.1.7 and 2.2.2.

### 3.3.2 Methods for Evaluating Variability and Uncertainty

***There are methodological differences regarding how variability and uncertainty are addressed in a Monte Carlo analysis. Two-dimensional simulations provide a means for distinguishing between variability and uncertainty in the overall Monte Carlo analysis, but there was a lack of consensus among the panelists on the appropriateness of a two-dimensional approach for practical applications.***

The statements prepared by Frey and Ferson (Appendix E) and Burmaster (Appendix E), as well as the case studies prepared by Price and Barry (Appendix D) and Teresa Bowers[6] provide insight into the methodological differences associated with addressing variability and uncertainty. Two-dimensional (2-D) simulations provide a formal approach to evaluating and distinguishing among the various components of uncertainty in the analysis. Other methods, such as interval analyses, have also been proposed. The panelists also discussed use of a tiered strategy for approaching uncertainty analysis (see Section 3.1.2).

Panelists discussion on this topic can be found in Section 2.2.2. Some of the methodological issues highlighted during this discussion are summarized below:

*Methodological Issues Related to Uncertainty*

- Standard data analysis tends to <u>understate uncertainty</u> by focusing solely on random error within a data set. Experience indicates that unsuspected systematic errors are usually substantial—requiring expansion of confidence limits.
- Various types of model errors represent important threats to the accuracy of uncertainty assessment. These include aggregation errors, common-mode failures, and uncertainties in the forms of quantitative relationships (e.g., dose-response functions). Distributional lumpiness can be caused by several types of model uncertainties.

---

[6]The case study presented by Teresa Bowers will be published as a paper authored by Cohen et al. in the December 1996 issue of *Human & Ecological Risk Assessment: An International Journal* under the title "The Use of Two-Stage Monte Carlo Simulation Techniques To Characterize Uncertainty and Variability."

*Methodological Issues Related to Variability*

- Standard data analysis tends to overstate variability by implicitly including measurement errors.
- Variability depends on the averaging time, averaging space, or other "dimensions" in which data are aggregated.
- A major threat to the accuracy of a data analysis is the representativeness of the population studied.
- Distributional lumpiness can be caused by cases where a small number of factors or a small number of discrete states are important.

***Methods should consider the stability of results at the tails of the cumulative distributions because this is where information is often used for decision-making.***

Decision-makers often rely on information at the tails to provide reasonable upper bounds of levels of confidence to support management decisions. There is a concern that Monte Carlo analytical results may be very uncertain or unstable at the tails of the distribution where information typically is sought by risk managers. In formulating the assessment, the risk assessor and the risk manager should discuss where in the distribution information will be most needed by the risk manager.

Workshop participants also suggested that sampling should be directed toward those parameters that are judged to be most important to the final result. In some cases, this might involve more intensive sampling at the tails of input distributions. When employing this method, the analyst should recognize that it is a stratified sampling scheme and that each of the samples has less probability than the overall probability.

Another issue concerning the tails of the distribution involves the quality of information at the tails of the input distributions, inasmuch as this is the information that most contributes to the tails of the output. Typically, the analyst has the least information about the input tails—both in terms of actual data and estimating uncertainty. This suggests two points:

- Data-gathering efforts should be structured to provide adequate coverage at the tails of the input distributions (especially at the bounds that yield higher exposure estimates), as well as at the more central values.

- The exposure assessment should include a narrative and qualitative discussion of the quality of information at the tails of the input distributions.

Workshop participants agreed that this area needs further attention. David Burmaster prepared a statement, reproduced in Appendix E, about approaches to ensuring the stability of Monte Carlo results at the tails.

***Alternative plausible conceptual or mathematical models are a potentially important source of uncertainty.***

Where important uncertainties are anticipated as a result of alternative plausible conceptual or mathematical models of exposure, it may be useful to conduct the Monte Carlo analyses separately for each alternative plausible model. This is analogous to a scenario-based approach (see Section 3.2.1), where the scenarios reflect different views or models of the exposure phenomenon.

Alternatively, the output from the different models could be treated as uncertain variables within a two-dimensional Monte Carlo computational framework. A difficulty in this latter approach is whether and how to weight the outputs of the alternative models.

Finally, from a practical or regulatory standpoint, the analysts may need to rely on one model. In such cases, the uncertainty associated with alternative plausible models is not addressed quantitatively. This should be recognized as a limitation of the uncertainty analysis. Discussion on this topic can be found in Section 2.2.2.

***There are limits to the analyst's ability to account for and characterize all sources of uncertainty. Where possible, the analyst should identify areas of uncertainty and consider how they might be included in the analysis, either qualitatively or quantitatively.***

Accounting for the important sources of uncertainty should be a key objective in Monte Carlo and other uncertainty analyses. However, it is not possible to characterize all the uncertainties associated with conceptual and/or analytical models. The analyst should attempt to identify the full range of types of uncertainty impinging on an analysis and clearly disclose what set of uncertainties the model attempts to represent and what it does not. In some cases, it may be prudent to use judgment to adjust distributions or parameters to account for unknowns, where these have been identified. Discussion on this topic can be found in Section 2.2.2.

### 3.3.3 The Role of Bayesian Methods

***Bayesian methods may be helpful for incorporating subjective information into uncertainty analyses in a manner that is consistent with distinguishing variability and uncertainty.***

Controversy continues to exist over the extent to which probability distributions are appropriately used to represent uncertainty in expert knowledge, versus their more traditional use in fitting data. While classical statistical methods for fitting distributions attempt to consider only the information contained in the data, Bayesian statistical methods explicitly allow for the incorporation of subjective, expert knowledge and judgment when developing distributions. However, because they allow the knowledge in the expert judgment to be combined with the information in the data, Bayesian methods have the capacity to bridge the gap between those who focus on expert knowledge in developing distributions and those who put greater emphasis on lab or field data.

Bayesian methods are quite compatible with efforts to separate variability and uncertainty in exposure and risk assessment. They also are very useful for designing experiments and data collection programs to reduce uncertainty. However, Bayesian methods can be computationally intensive and conceptually difficult to grasp for many. Therefore, exposure and risk assessors will require the assistance of a competent statistician experienced with Bayesian methods to apply these techniques to their applications. While this may preclude their application in many cases, increased use of Bayesian methods is likely in the future because of the types of problems they can address and the associated insights and benefits they provide. Mitchell Small prepared a

statement, reproduced in Appendix E, that elaborates on the use of Bayesian methods for Monte Carlo analysis.

## 3.4 PRESENTING RESULTS OF MONTE CARLO ANALYSIS

Panelists (most of whom have a technical orientation toward Monte Carlo analysis) noted that a full discussion about presenting results should include input from risk managers, the public, and others who are the primary audiences for these types of presentations. The principles articulated below lack that input, but may be useful as a starting point for further discussions. Panelists' discussions on presenting results can be found in Section 2.3.

***Presentations should be tailored to address the questions and information needs of the audience.***

Two primary audiences for the results of probabilistic risk assessment are the public and decision-makers. These audiences may have significantly different needs and interests regarding the presentation. Ideally, analysts should take time prior to presenting the results (preferably before even starting the analysis) to get to know the audience's needs and expectations regarding the results, as well as their overall knowledge about risk assessment in general and the Monte Carlo process in particular. This information will help analysts tailor their presentation to the audience. Where appropriate, involving a communications specialist in all phases of the assessment, beginning with problem formulation, may be helpful in improving communication between the scientists and the public and in targeting the presentation of results.

***Risk managers, site managers, and stakeholders should be informed and included to the extent possible in developing and implementing Monte Carlo analyses, rather than simply being called in to receive a final presentation of results. Stakeholder input may be accomplished via periodic meetings and briefings that address plans, methods, interim results, and possible outputs (both form and content).***

The information needs of the managers and stakeholders should be identified *throughout* the risk assessment management process to appropriately focus the analysis and to establish optimal methods for communicating results. Limiting contact between the analyst and risk managers and other stakeholders to a final presentation describing the analysis and results generally will be inadequate.

If a final public presentation is planned, the public should be involved in providing input to and commenting on the proposed analytical approach early in the process where possible. This type of involvement provides an opportunity for analysts to:

- Incorporate public needs and concern when designing the analytical framework.
- Build public trust regarding the analysis.
- Understand the audience's needs vis-a-vis presentation.
- Educate public stakeholders and decision-makers as the process evolves so that they can better understand the final results. This is particularly important because Monte Carlo analyses may involve terminology and yield results that are more complicated than those of a simpler exposure analysis.

As appropriate, progress on the analysis can also be discussed with stakeholders on an ongoing basis, culminating in a final presentation. This type of involvement will not always be possible, however.

***A tiered presentation style, in which briefing materials are assembled at various levels of detail, may be helpful.***

Entirely different types of reports are needed for scientific and nonscientific audiences. Scientists generally will want more detail than nonscientists. Risk managers may need more detail than the public. Reports for the scientific community are usually very detailed. Descriptive, less detailed summary presentations (e.g., box and whiskers, simple tables) and key statistics with their uncertainty intervals are generally more appropriate for nonscientists.

To handle the different levels of sophistication and detail needed for different audiences (or different segments within a single audience), it may be useful to design a presentation in a tiered format, where the level of detail increases with each successive tier. For example, the first tier could be a one-page summary that might include a graph or other numerical presentation as well as a brief description outlining what was done. This tier alone might be sufficient for some audiences. The next tier could be an executive summary, and the third tier could be a report. The second and third tiers would be useful for audiences, or portions of audiences, that want more detail than simply the "bottom line."

Within each major tier, information can be presented in stages. For example, rather than show a single complex chart, the information can be presented as a series of slides or overheads that begin with more general overview information and progress to more detailed information up to the level of detail needed by the target audience.

To learn about the information needs of EPA decision-makers, EPA's Office of Air Quality Planning and Standards convened a focus group of high-level EPA decision-makers across different offices and programs throughout the Agency (see Bloom et al. [1993], "Communicating Risk to Senior EPA Policy Makers: A Focus Group Study," pages 30-32, in Appendix F). The study showed that EPA risk managers need a variety of qualitative information, as well as quantitative risk measures, when making regulatory decisions or recommendations. "Recommendations for Presenting Monte Carlo Results to Risk Managers" in Appendix E summarizes a number of the findings and recommendations from the Bloom et al. study, which are incorporated into this recommendation by reference.

In addition, other institutions, such as the Congressional Research Service or other agencies that present information to high-level decision-makers might have some valuable presentation models that could be of use in designing presentations formats for risk information.

***A presentation should provide detailed information on the input distributions selected. This information should distinguish between uncertainty and variability.***

One of the higher, or more detailed, tiers of a results presentation should include detailed information on the input distributions selected. This detailed information should distinguish between variability and uncertainty and should include graphs and charts to visually convey written information.

Such information is important to thoroughly document and convey critical choices underlying the assessment that provide an important context for understanding and interpreting the results. Panelists agreed to incorporate, by reference, suggestions made by Burmaster and by Anderson (1994) (see Appendix F) and by Max Henrion and Tom McKone in their presentations (see Appendix D). These suggestions are summarized in "Recommendations for Presenting Information About Input Distributions" in Appendix E.

## 3.5 RECOMMENDATIONS

Workshop panelists offered the following recommendations:

***Develop a tiered exposure and uncertainty assessment strategy with guidance on how to move from one tier to the next.***

Some ideas on a tiered strategy are presented in "A Tiered Approach to Uncertainty/Variability Analysis in Exposure Assessment" in Appendix E.

***Develop guidance on how to develop site-specific or case-specific input distributions.***

The presentation by David Burmaster on "Input Data/Distributions for Model Parameters" (see Appendix D) includes information about the mechanics of deriving and using input data and may be a useful starting point for this type of guidance.

*Compile and develop a library of initial input distributions.*

These distributions would provide analysts with a common starting point for their specific exposure assessments. They should include supporting information regarding their derivation, use, and limitations. They could serve to complement EPA's Exposure Factors Handbook. Several potential sources for initial distributions already exist, including those developed by the American Industrial Health Council and some states.

*Develop case studies on the successful application of Monte Carlo analysis in exposure assessment for use by analysts and reviewers.*

The case studies should cover a range of applications.

*Develop criteria for reviewing and evaluating the quality of Monte Carlo analyses* (i.e., how to tell a "good" analysis from a "bad" one).

- A workgroup chaired by Dale Hattis and Paul Price prepared an initial list of suggested criteria (see Appendix E) that can serve as a starting point for the more detailed guidance needed.
- In addition, a Burmaster and Anderson (1994) article on "Principles of Good Practice for the Use of Monte Carlo Techniques in Human Health and Ecological Risk Assessments" (see Appendix F) also provides useful material for future guidance in this area. The paper outlines 14 principles of good practice that a reviewer can use to help judge the quality and content of Monte Carlo reports.
- *Communicating and Documenting Uncertainty in Risk Analysis* in Appendix D also provides useful material.

*Develop guidance for reports of Monte Carlo analyses.*

The guidance should define a report structure and format that provides for clear and logical presentation of results.

***Develop guidance on communicating and presenting Monte Carlo approaches and results.***

To develop this guidance, it may be useful to convene focus groups of risk managers and technical analysts similar to the focus groups convened for the Bloom et al. (1993) study (see Appendix F). At a minimum, further discussion is needed between managers and technical analysts regarding how Monte Carlo analysis can be used as a tool to meet management goals and objectives using Monte Carlo. Analysts need insight from risk managers about the kind and quality of information needed for decision-making.

# APPENDIX A

# DISCUSSION ISSUES

# Workshop on Monte Carlo Analysis

## U.S. Environmental Protection Agency
## New York, NY
## May 14-16, 1996

### Discussion Issues

This workshop is being held to discuss general principles for the use of Monte Carlo analysis in exposure assessment for human health risk assessment. The workshop discussions will focus on the technical issues concerning how to perform the analysis. Although these technical issues play a role in determining when to apply Monte Carlo techniques, the question of when is policy oriented and involves time and resource considerations. Policy issues concerning when to use Monte Carlo techniques will *not* be a focus of this May workshop.

The basic steps in developing an exposure assessment include:

- clearly defining the assessment questions and needs; developing a conceptual model which addresses these questions;
- selecting or deriving a mathematical model;
- identifying and selecting data for the model input parameters;
- evaluating the variability and uncertainty in the input parameters and their effect on the variability and uncertainty in the model output; and
- presenting the results.

Clearly defining the assessment questions and developing the conceptual model result from a full understanding of the information needs of the risk manager. Case specific issues—such as the size of the population of concern and the need to consider various subpopulations—are factored together with more generic issues, such as statutory requirements and acceptable health criteria, to define the assessment questions. Once defined, the assessment questions "drive" the remainder of the exposure assessment process. Selection or derivation of a mathematical model follows from the development of the conceptual model. In addition, technical issues such as the model's ability to represent the temporal nature of the exposure of the population of interest must be considered (i.e., short term, intermediate term, and/or chronic). Inherent in this is the model's ability to account for the dynamic behavior of the chemical of concern, the exposure media, and the receptor individuals within the population.

Many of the technical issues that arise during the application or review of Monte Carlo analyses occur during the last three steps in the exposure assessment process:

- selecting input data/distributions for model parameters;
- evaluating variability and uncertainty; and
- presenting results.

These issues will be the focal points for developing principles during this workshop. Case studies and papers will be used to highlight the issues and explore approaches for solving the problems. Ideally, the full set of case studies used for the workshop should cover all of these issues. Further, it would be an advantage if multiple case studies highlight the same issue and offer different approaches for resolution.

## Input Data/Distributions for Model Parameters

1. It has been suggested that prior to performing a Monte Carlo analysis one should develop point estimates of exposure using traditional techniques. Then, a sensitivity analysis is performed for each parameter in the exposure equation to determine which one/s have the most influence on the final result. It has been further suggested that the development and use of probability distributions be limited to those exposure parameters that have the most influence on the final result. Does this process represent the majority of expert opinions? How can one be confident that the sensitivity analysis, performed using traditional techniques, has identified important parameters for distributional analysis? How can one adequately characterize the uncertainty and variability in the output distribution when there is a mix of point estimates and probability distributions serving as input parameters?

2. For some parameters of the exposure equation, site-specific empirical measurements may not be available to determine the probability distributions. In these cases, distributions derived from surrogate data (e.g., national data on body weights) may be used. How do you characterize the uncertainty that has been introduced into the analysis when using surrogate data that are not collected from the population being studied? If surrogate data are inappropriate for evaluating exposure to the population of interest, site-specific empirical measurements may be necessary. What guidance can be given on the collection of site-specific empirical data that are collected to replace the surrogate data used to develop the distribution for a particular exposure parameter? How can you handle subpopulations when developing these data? How can you characterize the reduction in uncertainty associated with the collection of the new data?

3. In some cases, empirical measurements (site-specific or otherwise) for a particular exposure parameter may not be available or may be inadequate to determine a probability distribution. In these situations, should a distribution be estimated to complete the Monte Carlo analysis? If so, how? For example, it has been proposed that distributions for these parameters may be estimated via expert judgement or Delphi techniques. If these techniques are used, what factors should be considered in the weight of evidence? Can the effect (of using a distribution derived in this manner) on the tails of the output distribution be characterized? If so, how?

4. Some of the parameters in the exposure calculation may be correlated with each other. Which parameters do we presently know are correlated? Do we know the magnitude of the correlations that exist? These correlations may vary in strength and the absolute value of the correlations are often unquantified/unquantifiable. If these correlations exist and are moderate to strong, they may have effects on the tails of the output distributions. How should these correlations be accounted for in the Monte Carlo analysis? For example, it has been proposed that one may perform one Monte Carlo simulation with the correlations set to zero and another with the correlations set to some plausibly high value. In this way, the analyst may evaluate the importance of unquantified correlations in the analysis.

5. Empirical data collected from short-term studies of a particular exposure parameter may be inappropriate for use in Monte Carlo analysis when evaluating chronic exposures. For which exposure parameters are extrapolations from short-term data to chronic exposure appropriate? For which are extrapolations inappropriate?

## Evaluating Variability and Uncertainty

1. How can one adequately characterize the uncertainty associated with the selected conceptual and mathematical models? Can all types of variability and uncertainty be analyzed using techniques such Monte Carlo analysis?

2. Distributions of commonly used exposure parameters may reflect: uncertainty alone, variability alone, uncertainty and variability together, or variability and some restricted or biased measure of uncertainty. How can one be confident that the input distributions capture and represent both the variability and the uncertainty in the input exposure parameters?

3. How can one adequately characterize the uncertainty associated with Monte Carlo output distributions (e.g., developing confidence intervals around projected exposure estimates on the distribution curve?).

4. It has been suggested that keeping variability and uncertainty separate throughout a probabilistic assessment is essential. Further, it has been suggested that variability and uncertainty are best dealt with simultaneously within a Monte Carlo analysis by a 2-dimensional analysis. Do these suggestions represent the majority of expert opinions? Do these suggestions apply in all Monte Carlo analyses? What other approaches are there for dealing with variability and uncertainty within a Monte Carlo analysis?

5. How can one evaluate the numerical stability of the tails of the output distribution? It has been suggested that performing greater than or equal to 10,000 iterations in the Monte Carlo simulation and using software that includes Latin Hypercube Sampling will help to stabilize the tails. Does this suggestion represent the majority of expert opinions and does it apply in all situations? What other approaches are there for stabilizing the tails of the output distributions?

## Presenting Results

1. What are the basic elements that must be present in a report that presents the results of a Monte Carlo analysis? How can the results of the analysis be checked for quality assurance purposes (e.g., reproduce 10 percent of the calculations using software different from that used in the analysis under review).

2. How can the variability and uncertainty in the analysis be adequately characterized and discussed in the presentation of results so as not to overstate the precision of the analysis?

3. What is the best way to compare point estimates of exposure to the output of the Monte Carlo simulation? What information should be included in the discussion of this comparison? How can the benefits and limitations of Monte Carlo analysis over the point estimate technique be adequately characterized? What discussion is needed when the results of the Monte Carlo simulation are significantly different than the point estimates?

4. If a sensitivity analysis has been conducted, how can one adequately characterize and discuss the results? How can one characterize the influence of the sensitivity analysis on the selection of point estimates or probability distributions for input parameters?

5. How can one adequately characterize the stability of the tails of the output distribution? Can an adequate discussion of the confidence in the high-end values be provided? For example, what confidence does the analyst have that the high-end value is a realistic, but low probability, event. Can the possibility that the combinations of exposure parameters in the Monte Carlo simulation may result in an estimate of exposure which greatly exceeds the true value be adequately addressed?

6. When professional judgement or Delphi techniques are used to estimate distributions for input parameters, what is the best way to describe the process in the presentation of results? Which factors weighing into the decision should be listed? How can the potential effect on the output distribution be characterized?

7. The output distribution of exposure estimates may be incompatible with the dose-response endpoint selected for quantitative risk assessment because of fixed exposure assumptions imbedded in the toxicity metrics. How can this situation be avoided? Should probability distributions be determined for the dose-response values?

# APPENDIX B

# LIST OF PANEL MEMBERS AND OBSERVERS

# Workshop on Monte Carlo Analysis

## U.S. Environmental Protection Agency
## New York, NY
## May 14-16, 1996

# List of Panel Members

**Elmer Akin**
*Chief, Office of Health Assessment*
Waste Management Division
U.S. Environmental Protection Agency
345 Courtland Street, NE
Atlanta, GA 30365
404-347-1586 Ext. 6361
Fax: 404-347-1918

**Timothy Barry**
Office of Policy, Planning,
and Evaluation
U.S. Environmental Protection Agency
401 M Street, SW (2137)
Washington, DC 20460
202-260-2038
Fax: 202-260-1935
E-mail: barry.timothy@epamail.epa.gov

**David Bennett**
Senior Process Manager for Risk
Office of Emergency
and Remedial Response
U.S. Environmental Protection Agency
401 M Street, SW (5202G)
Washington, DC 20460
703-603-8759
Fax: 703-603-9133
E-mail: bennett.da@epamail.epa.gov

**Teresa Bowers**
*Principal*
Gradient Corporation
44 Brattle Street
Cambridge, MA 02138
617-576-1555
Fax: 617-864-8469
E-mail: tbowers@gradcorp.com

**David Burmaster**
*President*
Alceon Corporation
P.O. Box 382669
Cambridge, MA 02338-2669
617-864-4300 Ext. 222
Fax: 617-864-9954
E-mail: deb@alceon.com

**Michael Callahan**
National Center for
Environmental Assessment
Office of Research and Development
U.S. Environmental Protection Agency
401 M Street, SW (8602)
Washington, DC 20460
202-260-8909
Fax: 202-401-1722

**Clark Carrington**
*Pharmacologist*
U.S. Food and Drug Administration
200 C Street, SW (HFS-308)
Washington, DC 20204
202-205-8705
Fax: 202-260-0498
E-mail: mzb@vm.cfsan.fda.gov

**Hsieng-Ye Chang**
*Environmental Engineer*
U.S. Army
ATTN: CMDR.USACHPPM
5158 Black Hawk Road
(MCHB-DC-EHR)
Aberdeen Proving Ground, MD 21010
410-671-2025
Fax: 410-671-8170

**Elizabeth Doyle**
*Toxicologist*
Health Effects Division
Office of Pesticide Programs
U.S. Environmental Protection Agency
401 M Street, SW (7509C)
Washington, DC 20460
703-308-2722
Fax: 703-305-5453
E-mail: elizabeth.doyle@epamail.epa.gov

**Michael Dusetzina**
*Environmental Engineer*
Office of Air Quality
Planning and Standards
Risk and Assessment Exposure Group
U.S. Environmental Protection Agency
(MD-15)
Research Triangle Park, NC 27711
919-541-5338
Fax: 919-541-0840

**Scott Ferson**
*Senior Scientist*
Applied Biomathematics
100 North Country Road
Setauket, NY 11733
516-751-4350
Fax: 516-751-3435
E-mail: risk@life.bio.sunysb.edu

**H. Christopher Frey**
*Assistant Professor*
Department of Civil Engineering
North Carolina State University
P.O. Box 7908
Raleigh, NC 27695-7908
919-515-1155
Fax: 919-515-7908
E-mail: frey@eos.ncsu.edu

**Susan Griffin**
*Toxicologist*
U.S. Environmental Protection Agency
999 18th Street (8EPR-PS)
Suite 500
Denver, CO 80202-2466
303-312-6562
Fax: 303-312-6065

**Annette Guiseppi-Elie**
*Environmental Scientist*
Exxon Biomedical Sciences, Inc.
Mettlers Road (CN-2350)
East Millstone, NJ 08875-2350
908-873-6730
Fax: 908-873-6009
E-mail: annette.guiseppi@exxon.sprint.com

**P.J. (Bert) Hakkinen**
*Senior Scientist, Toxicology and Risk Assessment*
Corporate, Professional, and Regulatory Services
The Procter & Gamble Company
Ivorydale Technical Center
5299 Spring Grove Avenue
Cincinnati, OH 45217-1087
513-627-6895
Fax: 513-627-4665
E-mail: hakkinenpj@pg.com

**Dale Hattis**
*Research Associate Professor*
Center for Technology, Environment, and Development
Clark University
950 Main Street
Worcester, MA 01610
508-751-4603
Fax: 508-751-4600
E-mail: dhattis@vax.clarku.edu

**Max Henrion**
*President*
Lumina Decision Systems, Inc.
4984 El Camino Road - Suite 105
Los Altos, CA 94002
415-254-0189 Ext. 27
Fax: 415-254-0292
E-mail: henrion@lumina.com

**Melanie Marty**
*Chief, Air Risk Assessment Unit*
Health Hazard Assessment
California Environmental Protection Agency
2151 Berkeley Way - Annex 11
Berkeley, CA 94704
510-540-3081
Fax: 510-540-2923
E-mail: mmarty@hw1.cahwnet.gov

**Thomas McKone**
*Adjunct Professor and Staff Scientist*
University of California
School of Public Health
Lawrence Berkeley Laboratory
140 Warren Hall #7630
Berkeley, CA 94720-7630
510-642-8771
Fax: 510-642-5815
E-mail: temckone@lbl.gov

**Charlie Menzie**
*Workshop Chair*
Menzie-Cura and Associates
One Courthouse Lane - Suite 2
Chelmsford, MA 01824
508-453-4300
Fax: 508-453-7260

**Marian Olsen**
*Environmental Scientist*
Technical Support Section
Program Support Branch
Emergency and Remedial Response Division
U.S. Environmental Protection Agency
290 Broadway
New York, NY 10007
212-637-4313
Fax: 212-637-4360

**Paul Price**
McLaren/Hart ChemRisk
Stroudwater Crossing
1685 Congress Street
Portland, ME 04102
207-774-0012
Fax: 207-774-8263
E-mail: paul_price@mclaren-hart.com

**Frank Schnell**
Agency for Toxic Substances and Disease Registry
1600 Clifton Road (E32)
Atlanta, GA 30333
404-639-0618
Fax: 404-639-0654
E-mail: fys0@atsdha1.em.cdc.gov

**Mitchell Small**
*Professor*
Departments of Civil & Environmental Engineering and Engineering & Public Policy
Carnegie Mellon University
Frew Street, Porter Hall
Pittsburgh, PA 15213-3890
412-268-8782
Fax: 412-268-7813
E-mail: ms35@andrew.cmu.edu

**Alan Stern**
*Acting Chief, Bureau for Risk Analysis*
Division of Science and Research
New Jersey Department of Environmental Protection and Energy
401 East State Street (CN-409)
Trenton, NJ 08625-0409
609-633-2374
Fax: 609-292-7340
E-mail: astern@dep.state.nj.us

**Daniel Wartenberg**
*Associate Professor*
Environmental and Occupational Health Sciences Institute
UMDNJ - R.W. Johnson Medical School
681 Frelinghuysen Road
P.O. Box 1179
Piscataway, NJ 08855-1179
908-445-0197
Fax: 908-445-0784
E-mail: dew@eohsi.rutgers.edu

**Paul White**
*Environmental Health Scientist*
National Center for Environmental Assessment
Office of Research and Development
U.S. Environmental Protection Agency
401 M Street, SW (8623)
Washington, DC 20460
202-260-2589
Fax: 202-260-3803

**Jeffrey Wong**
*Chief, Office of Scientific Affairs*
Department of Toxic Substances Control
California Environmental Protection Agency
400 P Street - 4th Floor
Sacramento, CA 95812-0806
E-mail: jjwngphd@netcom.com
or jffwng@ucdavis.edu

# Workshop on Monte Carlo Analysis

## U.S. Environmental Protection Agency
## New York, NY
## May 14-16, 1996

# Final Observer List

**Mary Ballew**
Environmental Scientist
U.S. Environmental Protection Agency
JFK Federal Building (HBT)
One Congress Street
Boston, MA 02203
617-573-5718
Fax: 617-573-9662

**Dennis Barkas**
Senior Hydrogeologist
Anderson, Mulholland
& Associates, Inc.
611 Broadway - Suite 907G
New York, NY 10012
212-505-9553
Fax: 212-505-9567

**Bob Benson**
Toxicologist
Division of Water
Municipal Facilities
U.S. Environmental Protection Agency
999 18th Street - Suite 500
Denver, CO 80439
303-312-7070
Fax: 303-312-6131

**John Blankinship**
Consulting Scientist
133 15th Street, NW
Washington, DC 20005

**Ruth Bleyler**
Environmental Scientist
U.S. Environmental Protection Agency
JFK Federal Building (HBS)
One Congress Street
Boston, MA 02203
617-573-5792
Fax: 617-573-9662

**Anne Marie Burke**
Regional Human Health Risk
Assessment Expert
U.S. Environmental Protection Agency
JFK Federal Building (HBS)
One Congress Street
Boston, MA 02203
617-223-5528
Fax: 617-573-9662

**Hsieng-Ye Chang**
Environmental Engineer
ATTN: CMDR.USACHPPM
U.S. Army
5158 Black Hawk Road
(MCHB-DC-EHR)
Aberdeen Proving Ground, MD 21010
410-671-2025
Fax: 410-671-8170

**Steven Chang**
Environmental Engineer
U.S. Environmental Protection Agency
401 M Street, SW (5204G)
Washington, DC 20460
703-603-9017
Fax: 703-603-9103

**James Cogliano**
Chief of Quantitative
Risk Methods Group
National Center for
Environmental Assessment
U.S. Environmental Protection Agency
401 M Street, SW
Washington, DC 20460
202-260-3814
Fax: 202-260-3803

**Fred Cornell**
Risk Analyst
Environmental Liability Management, Inc.
218 Wall Street
Research Park
Princeton, NJ 08540
609-683-4848
Fax: 609-683-0129
E-mail: elm@ix.netcom.com

**David Cozzie**
Regulatory Impact Analyst
U.S. Environmental Protection Agency
401 M Street (5307)
Washington, DC 20460
202-260-4294
Fax: 202-260-0284
E-mail: cozzie.david@epamail.epa.gov

**David Craigin**
Senior Toxicologist
Elf Atlchem
2000 Market Street
Philadelphia, PA 19103
215-419-5880
Fax: 215-419-5800

**Linda Cullen**
Unit Supervisor
New Jersey Department of
Environmental Protection and Energy
401 East State Street (CN-413)
Trenton, NJ 08625
609-984-9778
Fax: 609-292-0848

**Tod DeLong**
Senior Project Scientist
Life Systems Department
Roy F. Weston, Inc.
1 Weston Way
West Chester, PA 19380-1499
610-701-7304
Fax: 610-701-7401

**Susan Dempsey**
Nebraska Department of Health
301 Centennial Mall South
Lincoln, NE 68509
402-471-2541
Fax: 402-471-6436

**Janine Dinan**
Environmental Health Scientist
U.S. Environmental Protection Agency
401 M Street, SW (5202G)
Washington, DC 20460
703-603-8824
Fax: 703-603-9133

**Gina Ferreira**
Environmental Scientist
Emergency and Remedial
Response Division
U.S. Environmental Protection Agency
290 Broadway
New York, NY 10007
212-637-4431
Fax: 212-637-4360

**Sarah Foster**
Managing Associate
The Weinberg Group, Inc.
1220 19th Street, NW - Suite 300
Washington, DC 20036
202-833-8077
Fax: 202-833-4157

**David Gallegos**
Sandia National Laboratory
P.O. Box 5800 (MS-1345)
Albuquerque, NM 87185
505-845-0760
Fax: 505-848-0764

**Philip Goodrum**
Scientist
Syracuse Research Corporation
Merrill Lane
Syracuse, NY 13210
315-426-3429

**Joseph Greenblott**
Office of Research and Development
U.S. Environmental Protection Agency
401 M Street, SW (8104)
Washington, DC 20460
202-260-0467
Fax: 202-260-6932
E-mail: greenglott.joseph@epamail.epa.gov

**Peter Grevatt**
Environmental Scientist
Emergency and Remedial
Response Division
U.S. Environmental Protection Agency
290 Broadway
New York, NY 10007
212-637-4312
Fax: 212-637-4360

**Nicholas Gudka**
Vice President
Sciences International
King Street Station
1800 Diagonal Road - Suite 500
Alexandria, VA 22315
703-684-0123
Fax: 703-684-2223
E-mail: ngudka@sciences.com

**Kim Hoang**
ORD Regional Scientist
Office of Regional Administrator
U.S. Environmental Protection Agency
290 Broadway
New York, NY 10007
212-637-3591

**Karen Hogan**
Statistician
U.S. Environmental Protection Agency
401 M Street, SW (7403)
Washington, DC 20460
202-260-3895
Fax: 202-260-1279
E-mail: hogan.karen@epamail.epa.gov

**Ellen Ivens**
Senior Environmental Scientist
Anderson, Mulholland
& Associates, Inc.
611 Broadway
Suite 907G
New York, NY 10012
212-505-9553
Fax: 212-505-9567

**Betty Jensen**
Principal Consultant
Public Service Electric
and Gas Company
80 Park Plaza - 16th Floor
Newark, NJ 07101
201-430-6633
Fax: 201-504-8414
E-mail: bjensen@pseg.com

**Sheldon Jobe**
Senior Consultant
Booz-Allen & Hamilton, Inc.
8283 Greensboro Drive
McLean, VA 22102
703-917-2575
Fax: 703-917-3078
E-mail: jobeshel@bah.com

**Ashwin Kittur**
Toxicologist
CanTox U.S., Inc.
1011 Route 22 West
Bridgewater, NJ 08807
908-429-9202
Fax: 908-429-9260
E-mail: akittur@cntxmiss.mhs.compuserve.com

**Steven Knott**
Chemist
Risk Assessment Forum
U.S. Environmental Protection Agency
401 M Street, SW
Washington, DC 20460
202-260-1095
Fax: 202-260-3955

**Rao Kolluru**
Adjunct Professor, Risk
Assessment and Management
c/o CH2M Hill
Stevens Institute of Technology
99 Cherry Hill Road
Parsippany, NJ 07054
201-316-9300
Fax: 201-334-5847

**Arnold Kuzmack**
Senior Science Advisor
Office of Water
U.S. Environmental Protection Agency
401 M Street, SW (4301)
Washington, DC 20460
202-260-5821
Fax: 202-260-5394

**William Lowry**
Research Scientist
New Jersey Department
of Environment Protection
401 East State Street (CN-413)
Trenton, NJ 08625
609-633-1348
Fax: 609-292-0848

**Kurt Lunchick**
Exposure Assessment Consultant
Jellinek, Schwartz & Connolly, Inc.
1525 Wilson Boulevard - Suite 600
Arlington, VA 22209
703-312-8555
Fax: 703-527-5477
E-mail: 73414.252@compuserve.com

**Mark Maddaloni**
Environmental Scientist
Emergency and Remedial
Response Division
U.S. Environmental Protection Agency
290 Broadway
New York, NY 10007
212-637-4315
Fax: 212-637-4360

**Elizabeth Margosches**
Chief, Epidemeology and
Quantitative Methods Section
Health and Environmental
Review Division
Office of Pollution
Prevention and Toxics
U.S. Environmental Protection Agency
401 M Street, SW (7403)
Washington, DC 20460
202-260-1511
Fax: 202-260-1279
E-mail: margosches.elizabeth@epamail.epa.gov

**Alec McBride**
Environmental Protection Specialist
Office of Solid Waste
U.S. Environmental Protection Agency
401 M Street (5307)
Washington, DC 20460
202-260-4806
Fax: 202-260-0284
E-mail: mcbride.alexander@epamail.epa.gov

**Torin McCoy**
Environmental Toxicologist
Texas Natural Resource
Conservation Commission
P.O. Box 13087 (168)
Austin, TX 78711
512-239-1795
Fax: 512-239-1794

**Thomas McNevin**
Research Scientist
New Jersey Department of
Environmental Protection
401 East State Street (CN-413)
Trenton, NJ 08625
609-633-1348
Fax: 609-292-0848

**Jane Michaud**
Human Health Risk Assessor for
Federal Facilities Program
U.S. Environmental Protection Agency
JFK Federal Building (HBT)
One Congress Street
Boston, MA 02203
617-223-5528
Fax: 617-573-9662

**Kenneth Mitchell**
Risk Assessment Coordinator
Georgia Department of
Natural Resources
205 Butler Street, SE
Suite 1154
Atlanta, GA 30334
404-657-8645
Fax: 404-651-9425
E-mail: kenneth_mitchell@mail.dnr.state.ga.us

**Ligia Mora-Applegate**
Environmental Scientist
Florida Department of
Environmental Protection
2600 Blairstone road
Tallahassee, FL 32399
904-488-3935
Fax: 904-921-1815
E-mail: ligia@depstate.florida.us

**Sam Morris**
Head, Biomedical and
Environmental Assessment Group
Brookhaven National Laboratory
Building 490D
Upton, NY 11973
516-344-2018
Fax: 516-344-7867
E-mail: morris3@bnl.gov

**William Muszynski**
*Deputy Regional Administrator*
U.S. Environmental Protection Agency
290 Broadway - 26th Floor (2-RA)
New York, NY 10007
212-637-5000

**Joel O'Connor**
Ocean Policy Coordinator
U.S. Environmental Protection Agency
290 Broadway - 24th Floor
New York, NY 10007
212-637-3792
Fax: 212-637-3889

**Marian Olsen**
Environmental Scientist
Emergency and Remedial
Response Division
U.S. Environmental Protection Agency
290 Broadway - 18th Floor
New York, NY 10007
212-637-4313
Fax: 212-637-4360

**Amy Pelka**
Environmental Health Scientist
U.S. Environmental Protection Agency
77 West Jackson Boulevard (B-19J)
Chicago, IL 60604
312-886-9858
Fax: 312-353-5374

**Kimberly Pieslak**
Project Scientist
Blasland, Bouck, & Lee, Inc.
8 South River Road
Cranbury, NJ 08512
609-860-0590 Ext: 230
Fax: 609-860-8007

**Vince Pitruzzello**
Branch Chief
U.S. Environmental Protection Agency
290 Broadway
New York, NY 10007

**Lara Pullen**
Life Scientist
Water Division
ARTS Branch
U.S. Environmental Protection Agency
77 West Jackson Boulevard
Chicago, IL 60604
312-886-0138
E-mail: pullen.lara@epamail.epa.gov

**Hersch Rabitz**
Frick Chemical Laboratory
Princeton University
Washington Road
Princeton, NJ 08544
609-258-3917
Fax: 609-258-1595

**Lance Richman**
Remedial Project Manager
Emergency and Remedial
Response Division
U.S. Environmental Protection Agency
290 Broadway
New York, NY 10007
212-637-4409

**Ann Rychlenski**
Public Affairs Specialist
U.S. Environmental Protection Agency
290 Broadway - 26th Floor (2-EPD)
New York, NY 10007
212-637-3672
Fax: 512-637-4445

**Edward Sargent**
Director
Toxic and Environmental Health
Merck & Co., Inc.
P.O. Box 100 (WS2F-45)
Whitehorse, NJ 08889
908-423-7906
Fax: 908-735-1388
E-mail: edward_sargent@merck.com

**Brian Sassaman**
Bioenvironmental Engineer
U.S. Air Force
Armstrong Laboratores/OEMH
2402 E Drive
Brooks AFB, TX 78235
210-536-6122
Fax: 210-536-2315
E-mail: brian.sassaman@guardian.brooks.af.mil

**Stephen Schaible**
Biologist
Health Effects Division
Office of Pesticide Programs
U.S. Environmental Protection Agency
401 M Street, SW (7509C)
Washington, DC 20460
703-308-2470
Fax: 703-305-5142
E-mail: schaible.stephen@epamail.epa.gov

**Sophia Serda**
Regional Toxicologist
Hazardous Waste
Management Division
U.S. Environmental Protection Agency
75 Hawthorne Street (H-9-3)
San Francisco, CA 94105
415-744-2307
Fax: 415-744-1916
E-mail: serda.sophia@epamail.epa.gov

**Jeffrey Shorter**
Senior Research Scientist
Mission Research Corporation
1 Tara Boulevard - Suite 302
Nashua, NH 03062
603-891-0070 Ext: 215
Fax: 603-891-0088

**Smita Siddhanti**
Senior Associate
Booz-Allen & Hamilton, Inc.
8283 Greensboro Drive
McLean, VA 22102
703-917-2447
Fax: 703-917-3078
E-mail: siddhanti_smita@bah.com

**Elliot Sigal**
Scientist
CanTox U.S., Inc.
2233 Argentina Road - Suite 308
Mississauga, Ont L5N 2X7
905-542-2900
Fax: 905-542-1011
E-mail: esigal@spectranet.ca

**Anita Street**
Environmental Scientist
Office of Policy and Management
U.S. Environmental Protection Agency
290 Broadway
New York, NY 10007
212-637-3590

**Steven Su**
Bailey Research Associates
292 Madison Avenue
New York, NY 10017
212-686-1754
Fax: 212-685-6705

**Douglas Tomchak**
Remedial Project Manager
Emergency and Remedial
Response Division
U.S. Environmental Protection Agency
290 Broadway
New York, NY 10007
212-637-3956

**Paul Weathersby**
Automation Councilors, Inc.
5 Ferryview Drive
Gales Ferry, CT 06335
860-464-0304
Fax: 860-445-5806

**Michelle Wei**
Iowa Department of Public Health
Lucas State Office Building
Des Moines, IA 50319
515-281-8707
Fax: 515-242-6284

**William Wood**
Director
Risk Assessment Forum
Office of Research and Development
U.S. Environmental Protection Agency
401 M Street, SW (8103)
Washington, DC 20460
202-260-6743
Fax: 202-260-3955

# APPENDIX C

# AGENDA

United States
Environmental Protection Agency
Risk Assessment Forum

# Workshop on Monte Carlo Analysis

U.S. Environmental Protection Agency
New York, NY
May 14-16, 1996

# Agenda

**Workshop Chair:** **Charlie Menzie**
**Menzie-Cura & Associates, Chelmsford, MA**

**TUESDAY, MAY 14**

8:00AM **Registration/Check-In**

9:00AM **Welcome and Regional Perspective**
*William J. Muszynski, Deputy Regional Administrator, U.S. Environmental Protection Agency (U.S. EPA), New York, NY*

9:15AM **Overview**
*William Wood, U.S. EPA, Office of Research and Development (ORD), Risk Assessment Forum, Washington, DC*

9:30AM **Workshop Structure and Objectives**
*Charlie Menzie, Workshop Chair*

10:00AM BREAK

10:30AM **Topic Presentation: Input Data/Distributions for Model Parameters**
*David Burmaster, Alceon Corporation, Cambridge, MA*

11:15AM **Case Study Application: Benzene MACT**
*Michael Dusetzina, U.S. EPA, ORD, Washington, DC, and Charlie Menzie*

11:35AM **Case Study Application: Superfund Site**
*Teresa Bowers, Gradient Corporation, Cambridge, MA*

12:00PM LUNCH

1:30PM Panel Discussion

3:10PM BREAK

3:30PM
- Panel Discussion/Open Discussion
- Writing Assignments

5:00PM ADJOURN

## WEDNESDAY, MAY 15

8:30AM Planning and Logistics
*Charlie Menzie, Workshop Chair*

8:45AM **Topic Presentation: Variability/Uncertainty**
*Christopher Frey, North Carolina State University, Raleigh, NC*

9:15AM **Case Study Application: Radon in Drinking Water**
*Timothy Barry, U.S. EPA, Office of Policy, Planning, and Evaluation, Washington, DC*

9:35AM **Case Study Application: Superfund Site**
*Paul Price, McLaren/Hart ChemRisk, Portland, ME*

10:00AM BREAK

10:20AM Panel Discussion

12:00PM LUNCH

1:30PM
- Panel Discussion/Open Discussion
- Writing Assignments

2:00PM **Topic Presentation: Presenting Results**
*Thomas McKone, University of California, Berkeley, CA*

2:30PM Example(s) of Methods of Presenting Information to Decision-Makers and Risk Managers

3:00PM BREAK

3:20PM
- Panel Discussion/Open Discussion
- Writing Assignments

5:00PM ADJOURN

## THURSDAY, MAY 16

8:30AM **The General Principles: What Are the Main Points To Consider?**
*Charlie Menzie, Workshop Chair*

10:00AM BREAK

10:20AM **General Principles (continued)**

12:15PM Wrap-Up

12:30PM ADJOURN

# APPENDIX D

# WORKSHOP PRESENTATION MATERIALS

## INPUT DATA/DISTRIBUTIONS FOR MODEL PARAMETERS



## VARIABILITY/UNCERTAINTY



## PRESENTING RESULTS



---

[7]First appeared in: Price, P.S. et al. 1996. Uncertainty and variation in indirect exposure assessments: An analysis of exposure to tetrachlorodibenzo-*p*-dioxin from a beef consuption pathway. Risk Analysis 16(2):263-277. April. Reprinted with permission from Plenum Publishing Corporation.

# DEVELOPING INPUT DISTRIBUTIONS FOR PROBABILISTIC RISK ASSESSMENTS

**David Burmaster**
**Alcon Corporation**
**Cambridge, Massachusetts**

# Variability

Variability represents the natural heterogeneity or diversity in a well characterized *population*.

... is usually not reducible through further measurement or study.

... is a *bounded* characteristic or property of the *population*.

... is the primary physical, chemical, and biological phenomenon.

# Uncertainty

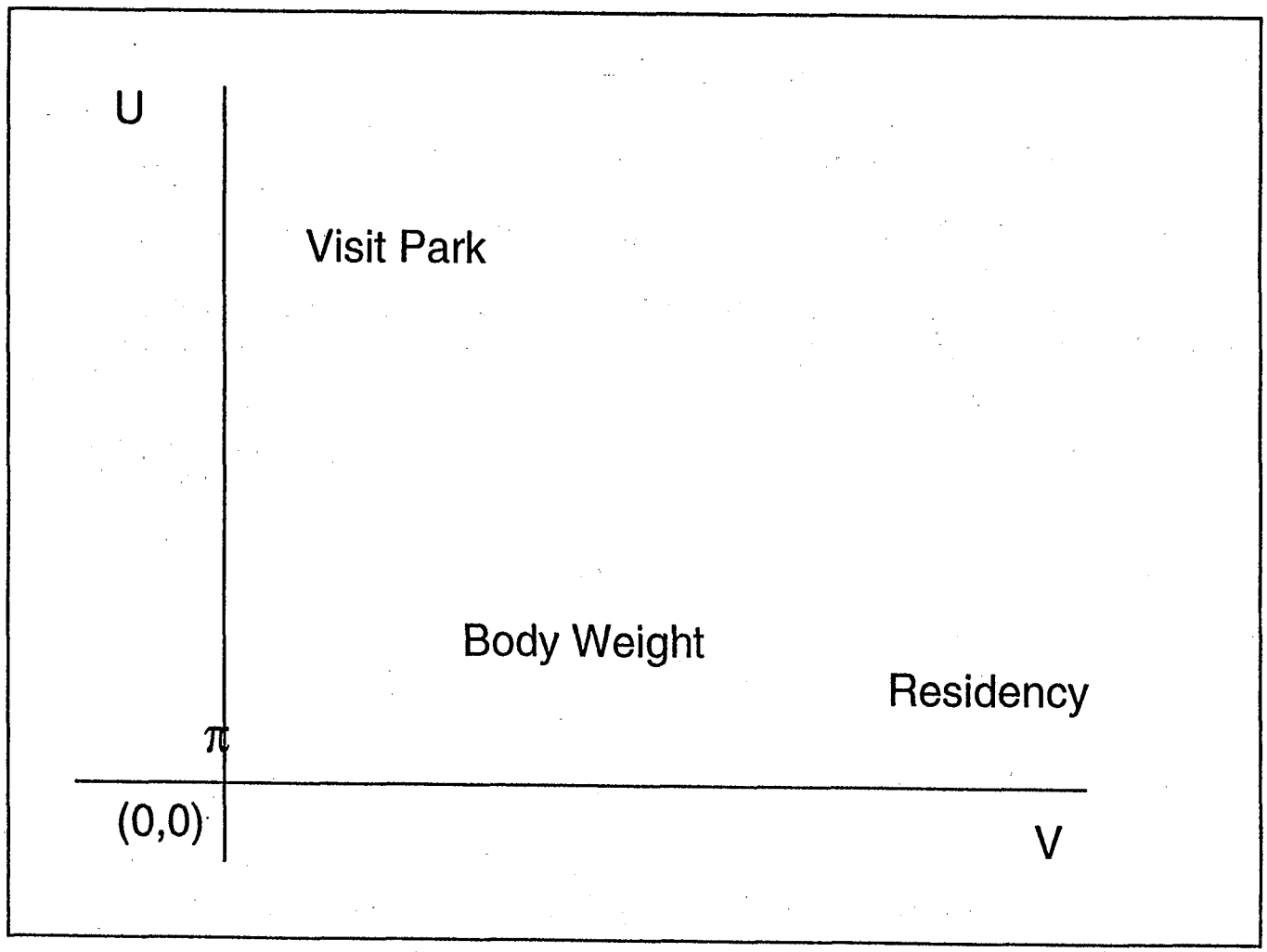
Uncertainty represents ignorance (or lack of perfect knowledge) about poorly-characterized phenomena or models.

... is sometimes reducible through further measurement or study.

... is an *unbounded* characteristic or property of the *analyst* .

... is the primary mental phenomenon.

# Relative Contributions



# Probabilistic Method, First-Order (1989):

$$\underline{\text{Risk}} = \frac{\prod \underline{X_i}}{\prod \underline{Y_j}} \bullet \underline{\text{CSF}}$$

where each variable is a (positive) first-order random variable (distribution) that encodes the variability and/or uncertainty

$$\underline{X} \sim \exp[\,\text{Normal}(\,\mu, \sigma\,)\,]$$

Here, V and U become intertwined

complications for risk assessor

complications for risk manager and public

# Probabilistic Method, Second-Order

$$\underline{\underline{\text{Risk}}} = \frac{\prod \underline{\underline{X_i}}}{\prod \underline{\underline{Y_j}}} \bullet \underline{\underline{\text{CSF}}}$$

where each variable is a (positive) second-order random variable that encodes both the variability and the uncertainty

$$\underline{\underline{X}} \sim \exp[\,\text{Normal}(\underline{\mu}, \underline{\sigma})\,]$$

# Which Variables to Make first-order RVs?

- All variables -- Or
- Select dominant variables by multiplying $E_i$ (Deterministic Framework)

$$E_i = \frac{\frac{\Delta R}{R}}{\frac{\Delta X_i}{X_i}} = \frac{\partial R}{\partial X_i} \bullet \frac{X_i}{R} = \pm 1 \text{ for simple models}$$

by $RSD_i$ (First-Order Probabilistic Framework)

$$RSD_i = CV_i = \frac{AStdDev(\underline{X_i})}{AMean(\underline{X_i})} \quad \text{or}$$

$$HighWidth_i = \frac{90th\ Percentile(\underline{X_i})}{Median(\underline{X_i})}$$

# Which Variables to Make second-order RVs?

- All variables -- Or
- Select dominant variables through computational experiments:
  - Sketch second-order RVs for all input variables
  - Run model with all variables as second-order RVs
  - Run model with group$_k$ variables "toggled" to simple RVs
  - Run model with all variables "toggled" to simple RVs
  - Run model group$_k$ variables "toggled" to second-order RVs
  - Watch the changes from ON <==> OFF to deduce dominant variables

# Processes Create Distributions

Physical Processes
erosion, fracture, accretion, dilution, ...

Chemical Processes
reaction, diffusion, ...

Biological and Toxicological Processes
susceptibility, enzyme variation, population (birth, life, death), ...

Statistical Processes
addition & subtraction, multiplication & division, exponentiation, ...

Mixture Processes
immigration, ...

Survival of a Cohort
pure death processes, replacement processes, ...

InterArrival Processes

# Presumptive Univariate Distributions

| | |
|---|---|
| Normal | Height |
| LogNormal | Body Weight, Skin Area, Inhalation, Diet, e.g., Fish, Drinking Water, Breast Feeding, Dust Transfer, Air Exchange Rate, House Volumes, Ambient Concentrations, Shower Duration, Soil Adherence, Lipid in Fish, PCB in Fish, more ... |
| Exponential & Weibull & Gompertz | Residency, Job Tenure |
| Poisson, Gamma, Exponential | InterArrival Times |
| Beta | Absorption Fraction |
| Uniform (Rectangular) | Professional Judgment |
| Triangular | Professional Judgment |

# Correlations and Functional Dependencies

Everything is correlated with everything, but ....

.... we need the data ....

Smith et al, 1992: $| \rho | < 0.6$ does not count in most circumstances

In human health risk assessment, only a few count,

SkinArea( BodyWeight) -->> a strong dependency

$\rho$ { ln(Body Weight), Height $\}_{adults}$ $\underline{\bullet}$ 0.3

Breathing(Metabolism, Activity, BodyWeight)

=====

but ...

Central Importance of Computational Experiments

# Fitting a Distribution to Data for V

Method of Moments
- easy, but no visualization and often abused

Probability Plots
- highly visual, but usually univariate

Maximum Likelihood
- high pedigree, but also need visualization

Maximum Entropy
- high pedigree, but often abused

# Fitting a Distribution to Opinion for U

Extrapolation

Professional Judgment

# Method of Moments

Normal( $\mu$, $\sigma$ )

$$\hat{\mu} = \text{AMean( data )} = \bar{x}$$

$$\hat{\sigma} = \text{AStdDev( data )} = s_x$$

LogNormal( $\mu$, $\sigma$ ) = exp[ Normal( $\mu$, $\sigma$ ) ]

$$\hat{\mu} = \text{AMean( ln ( data ) )} = \overline{\ln( x )}$$

$$\hat{\sigma} = \text{AStdDev( ln ( data ) )} = s_{\ln x}$$

Gamma( b, c )

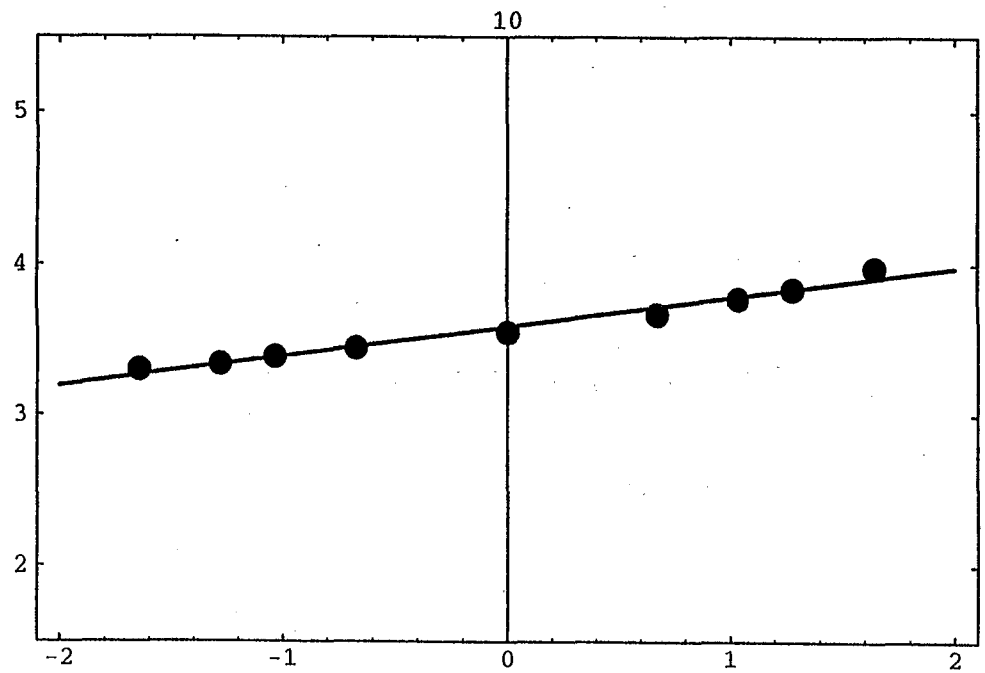
$$\hat{b} = \frac{(s_x)^2}{\bar{x}}$$

$$\hat{c} = \left( \frac{\bar{x}}{s_x} \right)^2$$

# Probability Plots

Highly visual, identify family of distribution and fit parameters

Normal, LogNormal, Exponential, others, but not Gamma



# Maximum Likelihood

Maximize the LogLikelihood function for model for a data set

Most powerful method for binned or nondetect data

$$J = \sum J_{pt} + \sum J_{nondet.pt} + \sum J_{bin.pt}$$

| | | |
|---|---|---|
| Jpt | = | ln( PDF(params, datum) |
| Jnondet.pt | = | ln( CDF(params, DL) ) |
| Jbin.pt | = | ln( CDF(params, top) - CDF(params, bot) ) |

Maximize J with respect to params -->> best fit

With "Profile Method," can obtain joint confidence regions

# Maximum Entropy

Profoundly important and often abused

Thermodynamics, Statistical Mechanics, Shannon's Information Theory

Signal Processing - extracting information from a noisy signal

- extract signal, discard noise
- remain most faithful to the evidence, introduce no artifacts



Support = [ - ∞ < min, max < + ∞ ]

| | | |
|---|---|---|
| no other knowledge | --->> | Uniform( min, max ) |
| mode | --->> | Triangular( min, mode, max ) |
| amean | --->> | truncated Exponential |
| gmean & gvar | --->> | truncated LogNormal |

Support = [ 0, + ∞ )

| | | |
|---|---|---|
| no other knowledge | --->> | no max entropy distribution |
| amean | --->> | Exponential |
| amean & gmean | --->> | Raleigh |
| gmean & gvar | --->> | LogNormal |

Uses in optical signal processing

Shaw & Tigg, 1994

# Delphi Techniques

Originated at RAND for defense studies (1960s)

Not standing around the water cooler or peering in a mirror

Do not lead the expert, even by suggesting a family of distributions

Often search for quartiles or percentiles, not parameters

Modern methodology

- Morgan & Henrion, 1990
- Cooke, 1991

# Selecting Data Sets

- start with national or state data

  NHANES, USDA, MI fish, FL seafood, ...

- general qualities

  random or stratified random sample

  large sample, long-term, different climates, different SES

  both genders, different ages, ...

  field-tested first

  disinterested respondents (and collectors)

# Steps

1. Find Good Data Set

2. View Data Using Exploratory Data Analysis

   Tukey
   Systat
   Cleveland
   Tufte

3. Analyze Data with as Few Assumptions as Possible

4. View
   Data
   Fit
   Residuals

5. Consider Mixtures, Go to 4.

6. Test Goodness of Fit

# Drinking Water Consumption

Data: In 1989, Ershow and Cantor published a statistical analysis of water intake rates for children and adults in different age groups as measured during and reported by the 1977-1978 Nationwide Food Consumption Survey of the USDA.

26K respondents
total water and tap water
age groups
- < 1 yr
- 1 - 11 yr
- 11 - 20 yr
- 20 - 65 yr
- > 65 yr

Information Reported:
12 bins, either 250 or 500 g/d wide

Method: Probability Plots for binned data

Subgroups: by age, could be done by region

Results: DW ~ exp[ Normal( $\mu$, $\sigma$) ]

For | z | $\leq$ 3.5 or 4, excellent fits

Reference: Roseberry & Burmaster, 1992

Limitations: 3-day survey, ... .

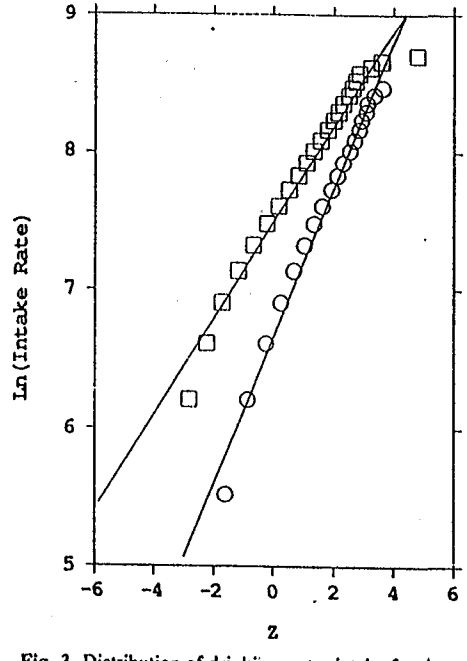



Fig. 3. Distribution of drinking water intake for the age group 11≤age<20: ln(intake rate) vs. $z$ (□ = total water; ○ = tap water).

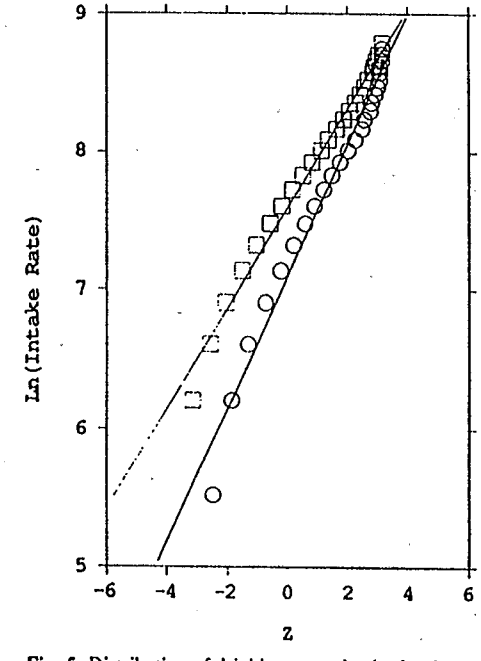



Fig. 5. Distribution of drinking water intake for the age group 65≤age: ln(intake rate) vs. $z$ (□ = total water; ○ = tap water).

# Drinking Water Consumption

Table II. Estimated Quantiles and Arithmetic Averages for Water Intake Rates

| Group | Total water intake (ml/d) Percentile 2.5 | 25 | 50 | 75 | 97.5 | Arithmetic average |
|---|---|---|---|---|---|---|
| 0 < age < 1 | 607 | 882 | 1074 | 1307 | 1900 | 1120 |
| 1 ≤ age < 11 | 676 | 1046 | 1316 | 1655 | 2562 | 1394 |
| 11 ≤ age < 20 | 907 | 1417 | 1790 | 2262 | 3534 | 1901 |
| 20 ≤ age < 65 | 879 | 1470 | 1926 | 2522 | 4218 | 2086 |
| 65 ≤ age | 970 | 1541 | 1965 | 2504 | 3978 | 2096 |
| All NFC survey | 807 | 1358 | 1785 | 2345 | 3947 | 1937 |
| *Simulated balanced population* | 808 | 1363 | 1794 | 2360 | 3983 | 1949 |
| | Tap water intake (ml/d) | | | | | |
| 0 < age < 1 | 80 | 176 | 267 | 404 | 891 | 323 |
| 1 ≤ age < 11 | 233 | 443 | 620 | 867 | 1644 | 701 |
| 11 ≤ age < 20 | 275 | 548 | 786 | 1128 | 2243 | 907 |
| 20 ≤ age < 65 | 430 | 807 | 1122 | 1561 | 2926 | 1265 |
| 65 ≤ age | 471 | 869 | 1198 | 1651 | 3044 | 1341 |
| All NFC survey | 341 | 674 | 963 | 1377 | 2721 | 1108 |
| Simulated balanced population | 310 | 649 | 957 | 1411 | 2954 | 1129 |

# Fish Ingestion

Data: In 1980, Rupp published an analysis of the one-year NMFS nationwide survey to provide a representative sample of fish consumption patterns among the population in Continental US.

- 23K participants
- 9 geographical regions of US
- 3 age groups: children, teenagers, adults
- 3 type of fish: salt water finfish, shellfish, fresh water finfish

Information Reported:

- median of daily consumption rate (DCR, in g/d)
- 90th percentile
- 99th percentile
- amean
- maximum
- count (n)

Method: Minimization of Objective Function, but
summary statistics instead of data points
many values reported as zeros

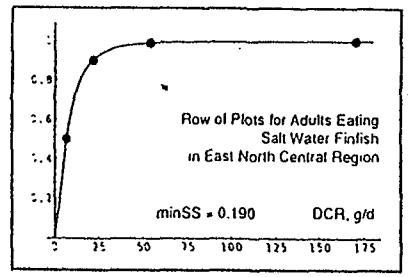
Finding: DCR ~ exp[ Normal( $\mu$, $\sigma$) ]

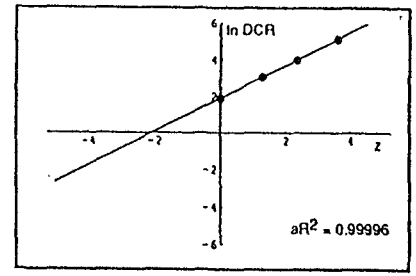
77 data sets were excellently fit by LogNormal distributions
12 data sets were adequately fit by LogNormal distributions
1 data set could not be fit by LogNormal distribution

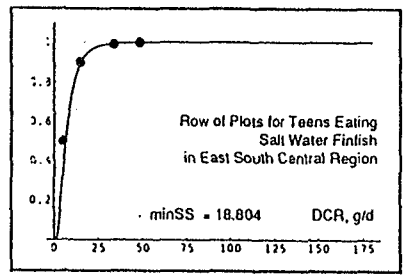
Reference: Ruffle et al, 1994

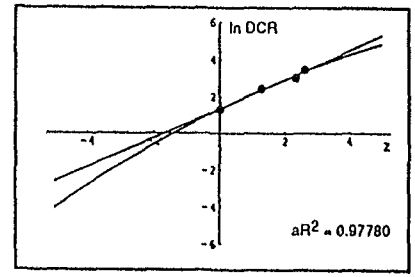
Limitations: old survey, not self-caught, ...

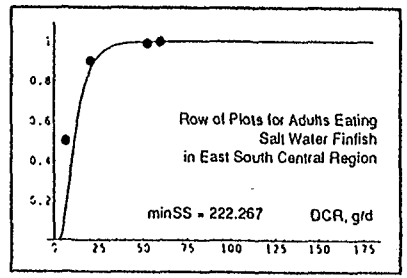
# Fish Ingestion

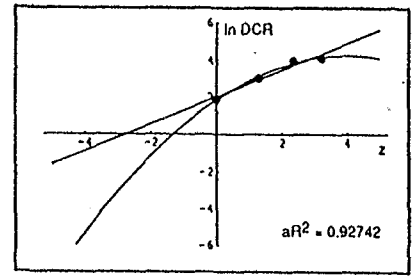



Column of Plots from NLO Method

Column of Plots from PP1 Method

# Height and Body Weight of Adults

Data: The US PHS conducted the NHANES II Survey from Feb 1976 through Feb 1980.

- 5.6K men and 6.5K women
- ages 18 - 74 yr
- statistically adjusted the raw data to reflect the whole US population:
  - age, sex, race

Information Reported

- Tables of Counts
  - 1-inch intervals in height
  - 10-lb intervals in weight

Method: Probability Plots and Minimization of Objective Function for binned data and statistical mixture of women

Subgroups: by age, all races and ethnicity, could be done by region

Results: $\{\underline{Ht}, \underline{BWt}\} \sim \text{Normal}(\mu_{Ht}, \sigma_{Ht}, \mu_{lnBWt}, \sigma_{lnBWt}, \rho)$

Men: Single statistical population
Women: Mixture of 2 statistical populations

Reference: Brainard & Burmaster, 1992

Limitations: old survey

Table I. Number of Men 18–74 Years of Age by Weight and Height, United States, 1976–1980 (Number of Persons in Thousands)[a,d]

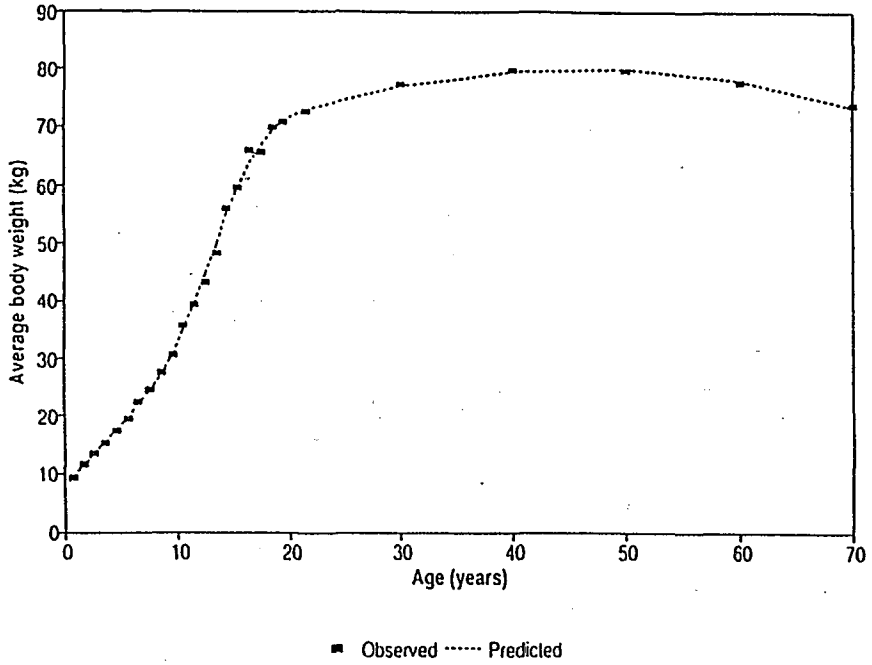
| Height[b] (in) | Weight[c] <110 (lb) | 110–119 (lb) | 120–129 (lb) | 130–139 (lb) | 140–149 (lb) | 150–159 (lb) | 160–169 (lb) | 170–179 (lb) | 180–189 (lb) | 190–199 (lb) | 200–209 (lb) | 210–219 (lb) | 220–229 (lb) | ≥230 (lb) | True total (lb) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| <62 | 41 | 70 | 100 | 42 | 110 | 38 | 69 | 24 | 8 | 19 | | 10 | 11 | | 542 |
| 62 | 38 | 34 | 94 | 102 | 196 | 73 | 35 | 33 | 48 | | 15 | | | | 668 |
| 63 | 66 | 65 | 195 | 197 | 286 | 136 | 113 | 98 | 33 | 29 | | 3 | | | 1221 |
| 64 | 33 | 110 | 237 | 381 | 376 | 413 | 181 | 231 | 106 | 62 | 7 | 8 | 30 | | 2175 |
| 65 | 53 | 191 | 177 | 578 | 806 | 820 | 556 | 363 | 269 | 161 | 154 | 30 | 30 | 25 | 4213 |
| 66 | 50 | 131 | 457 | 555 | 843 | 910 | 986 | 547 | 515 | 252 | 105 | 58 | 43 | 83 | 5535 |
| 67 | 12 | 102 | 324 | 780 | 1087 | 1237 | 1174 | 1181 | 801 | 429 | 319 | 135 | 154 | 245 | 7980 |
| 68 | 29 | 77 | 319 | 743 | 1127 | 1351 | 1625 | 1328 | 1152 | 686 | 390 | 284 | 250 | 205 | 9566 |
| 69 | 7 | 11 | 322 | 488 | 960 | 1169 | 1547 | 1436 | 1286 | 747 | 750 | 390 | 155 | 310 | 9578 |
| 70 | 4 | 37 | 104 | 455 | 900 | 1041 | 1450 | 1313 | 1334 | 710 | 479 | 441 | 252 | 347 | 8867 |
| 71 | | 32 | 22 | 242 | 453 | 911 | 818 | 1103 | 868 | 692 | 481 | 377 | 217 | 500 | 6716 |
| 72 | | 9 | 19 | 67 | 217 | 392 | 716 | 831 | 765 | 696 | 436 | 216 | 251 | 404 | 5019 |
| 73 | | | | 20 | 41 | 228 | 356 | 322 | 483 | 370 | 306 | 190 | 203 | 226 | 2745 |
| 74 | | | 7 | 42 | | 76 | 73 | 203 | 270 | 243 | 191 | 156 | 84 | 119 | 1464 |
| ≥75 | | | | | | 47 | 24 | 245 | 168 | 121 | 173 | 58 | 121 | 306 | 1263 |
| True total | 333 | 869 | 2377 | 4692 | 7402 | 8842 | 9723 | 9258 | 8106 | 5217 | 3806 | 2356 | 1801 | 2770 | 67,552 |

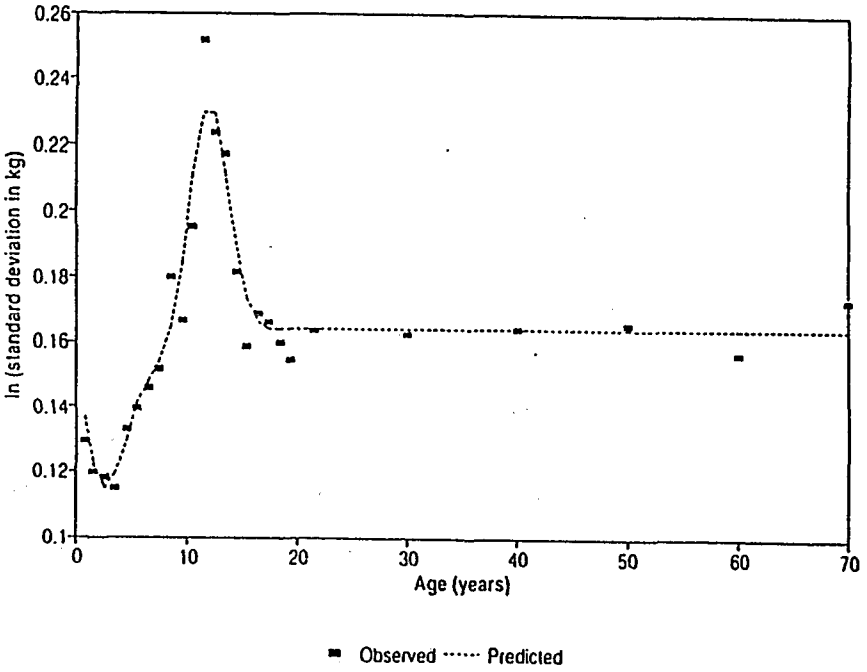
[a] *Source*: Ref. 4, Table 27.
[b] Height without shoes.
[c] Weight with clothes, estimated as ranging from 0.20–0.62 lb.
[d] Numbers in cells scaled up to reflect size of population; only 9983 men actually examined.



5 May 1996

# Body Weight as Function of Age



# Job Tenure

| | |
|---|---|
| Data: | US Bureau of the Census |
| Method: | Survival Analysis of complex mixture |
| Subgroups: | by gender, for several industries |
| Finding: | Tenure ~ Gompertz( duration ) |
| | Current Tenure ≠ Projected Tenure |
| Reference: | Shaw & Burmaster, 1995, in revision |
| Limitations: | general method, but only applied to a few industries so far |

# Job Tenure

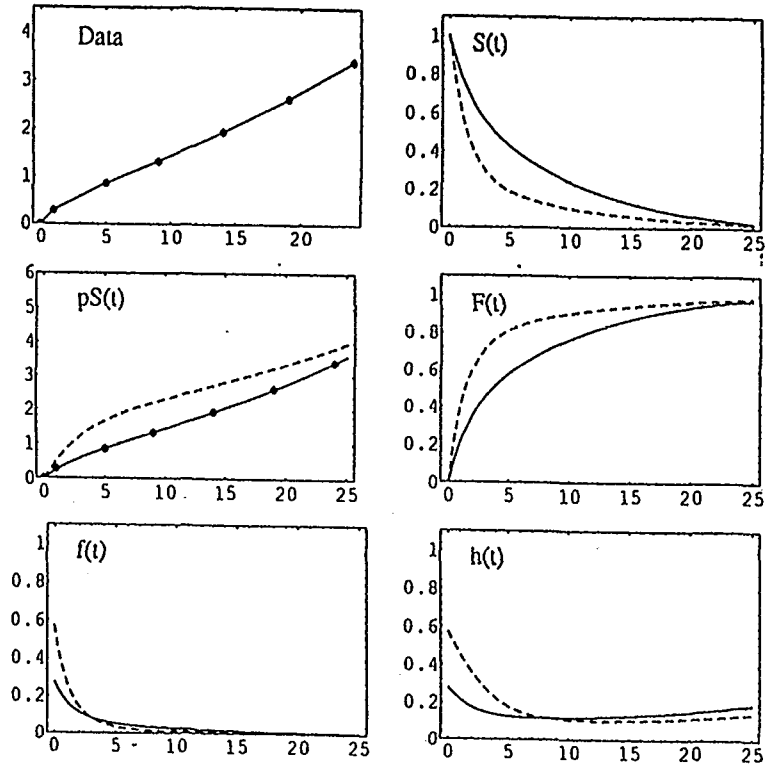



Figure 2
Women in Manufacturing
Plots of Data, S(t), pS(t), F(t), f(t), and h(t) for $t \in [0, 25]$ yr
with solid lines for Survey and dashed lines for Projection

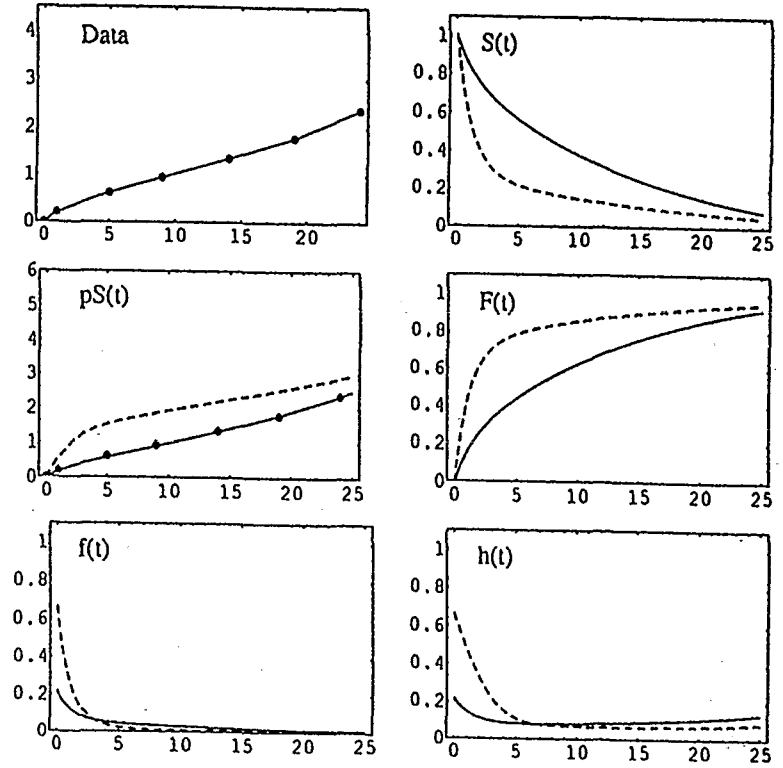



Figure 3
Men in Manufacturing
Plots of Data, S(t), pS(t), F(t), f(t), and h(t) for $t \in [0, 25]$ yr
with solid lines for Survey and dashed lines for Projection

# Truncation of Distributions

Natural Phenomena

Model

++++++++++

Need to capture the important features of the natural phenomena in the model

Need to prevent distortion of the natural phenomenon by the model

++++++++++

Some natural phenomena have lower and/or upper bounds

| | | | | |
|---|---|---|---|---|
| 0 | $\leq$ | days in week | $\leq$ | 7 |
| 0 | $\leq$ | fraction of skin | $\leq$ | 1 |
| 0 | $\leq$ | conc (ppb) | $\leq$ | $10^9$ |

Some do not

Some (parametric) distributions commonly used to model natural phenomena have lower and/or upper bounds

| | |
|---|---|
| Beta distribution | [ 0, 1 ] |
| Uniform distribution | [ min, max ] |
| Triangular distribution | [ min, max ] |

Some do not

| | |
|---|---|
| Normal distribution | ( $-\infty$, $+\infty$ ) |
| LogNormal distribution | [ 0, $+\infty$ ) |
| Exponential distribution | [ 0, $+\infty$ ) |

# My Personal Outlook on Truncation

Look at underlying process that generates the RV

If physical variable has an upperbound, truncate

Do not truncate at an arbitrary percentile

It takes a lot of information to know truncation

The principle of "least distortion" suggests ...

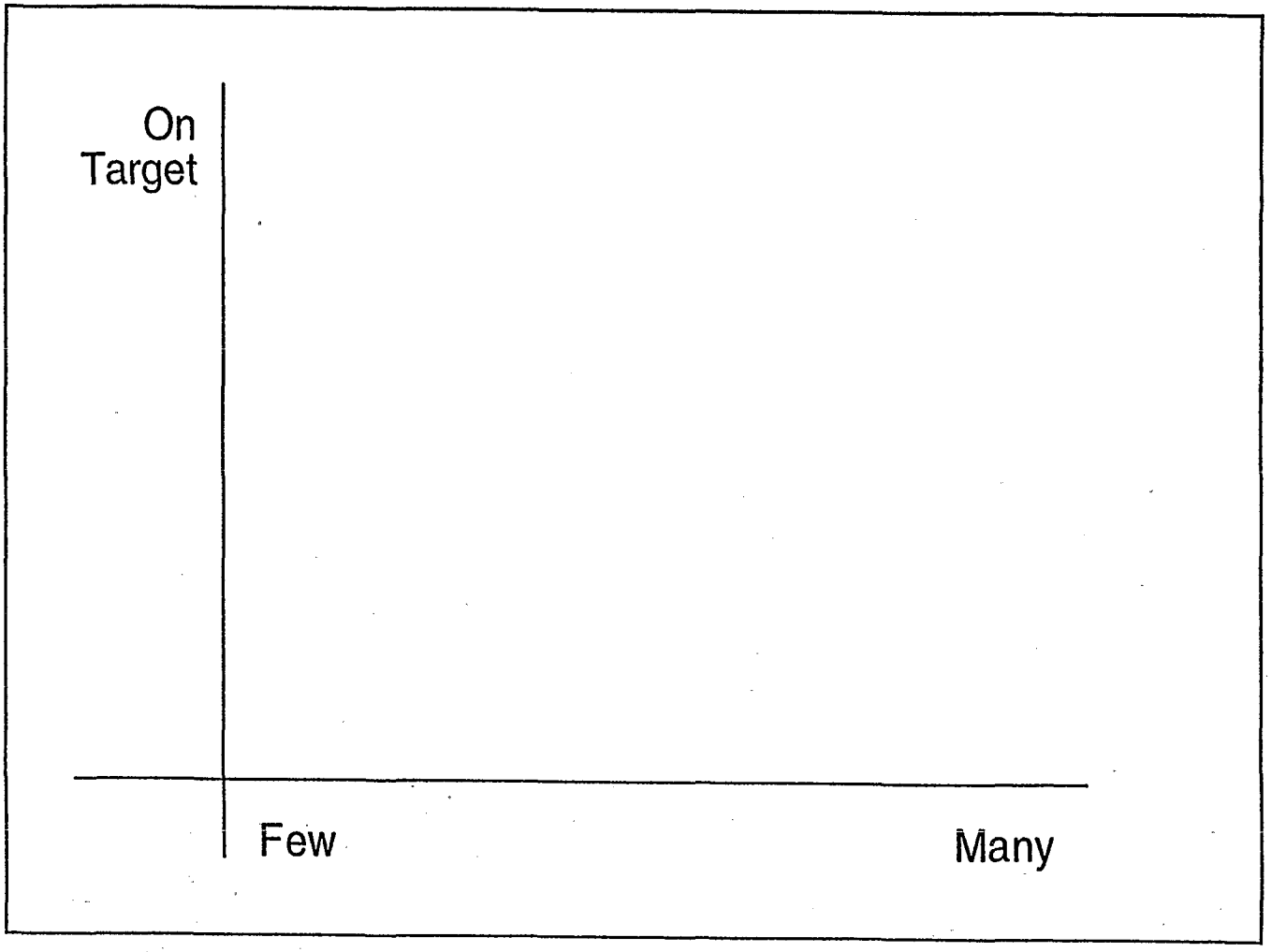
Central importance of computational experiments!!!!

RA must discuss results and issues for RM and public, including

- Value of additional information (aka, need for research)
- Sensitivity of findings to assumptions

Need to support and document choices made

# Extrapolation



# Extrapolation from Population to Population

Small to medium extrapolations by adjusting parameters, by keeping the family of the distribution

Self-caught fishing from ME to NH (fresh water)

House size from NY to PA (single-family, similar SES)

Large extrapolations by Delphi Method on parameters

Enormous extrapolations by Delphi Method on whole distribution

On whole distribution

==========

Uncertainty is a function of the analyst, not the population!!

# Extrapolation from Short to Long Duration

For whole distribution of single events, the tails only get wider ...

For E(x), the long-term average only gets tighter ...

===

NonStationarity can arise ...

Changes in diet

other ....

# Sensitive Subpopulations

Sensitivity based on physiology, biochemistry, and toxicology

- Pb causes more neurotoxicity to young children than to adults.
- Thalidomide causes damage to fetuses (during a window)

Sensitivity not based on exposure

- Number of years that a person lives in a house
  - Develop distribution for population in a town
  - Develop a distribution for owners in a town
  - Develop a distribution for owners who occupy >30 yr
  - Develop a distribution for owners who inherited house from parents ...

# Need for Continuing Innovation

- I hope that the first page of the Report from the Workshop will start with something along these lines:

> This Report contains guidelines for minimum practices that are acceptable for use in probabilistic exposure assessments.
>
> Given the breadth and depth of probabilistic methods, and given the rapid development of new probabilistic methods, we cannot list all the possible techniques that a risk assessor may use for a particular assessment.
>
> The US EPA emphatically encourages the development and application of new methods in exposure assessments, and nothing in this Report can or should be construed as limiting the development or application of new methods whose power and sophistication exceed the guidelines for minimum acceptable practice contained in this Report.

# References


AIHC, 1994
American Industrial Health Council, 1994, Exposure Factors Source Book, Washington, DC

Baird et al, 1996
Baird, S.J.S., J.T. Cohen, J.D. Graham, A.I. Shlyakhter, and J.S. Evans, 1996, Noncancer Risk Assessment: A Probabilistic Alternative to Current Practice, Human and Ecological Risk Assessment, Volume 2, Number 1, pp 79 - 102

Bogen, 1995
Bogen, K.T., 1995, Methods to Approximate Joint Uncertainty and Variability in Risk, Risk Analysis, Volume 15, Number 3, pp 411 - 419

Bogen, 1994
Bogen, K.T., 1994, A Note on Compounded Conservatism, Risk Analysis, Volume 14, Number 4, pp 379 - 381

Bogen, 1993
Bogen, K.T., 1993, An *Intermediate-Precision Approximation of the Inverse Cumulative Normal Distribution*, Communications in Statistics, Simulation and Computation, Volume 23, Number 3, pp 797 - 801

Bogen, 1992
Bogen, K.T., 1992, RiskQ: An Interactive Approach to Probability, Uncertainty, and Statistics for Use with Mathematica, Reference Manual, UCRL-MA-110232 Lawrence Livermore National Laboratory, University of California, Livermore, CA, July 1992

Bogen, 1990
Bogen, K.T., 1990, Uncertainty in Environmental Risk Assessment, Garland Publishing, New York, NY

Brainard & Burmaster, 1992
Brainard, J. and D.E. Burmaster, 1992, Bivariate Distributions for Height and Weight of Men and Women in the United States, Risk Analysis, 1992, Volume 12, Number 2, pp 267 - 275

Burmaster & Hull, 1996
Burmaster, D.E. and D.A. Hull, 1996, A Tutorial on LogNormal Distributions and LogNormal Probability Plots, in review

Burmaster & Anderson, 1994
Burmaster, D.E. and P.D. Anderson, 1994, Principles of Good Practice for the Use of Monte Carlo Techniques in Human Health and Ecological Risk Assessments, Risk Analysis, Volume 14, Number 4, pp 477 - 481

Burmaster, Lloyd & Crouch, 1994
Burmaster, D.E., K.J. Lloyd, and E.A.C. Crouch, 1994, LogNormal Distributions of Body Weight for Female and Male Children in the United States, Risk Analysis, in revision

Burmaster & von Stackelberg, 1991
Burmaster, D.E. and K. von Stackelberg, 1991, Using Monte Carlo Simulations in Public Health Risk Assessment: Estimating and Presenting Full Distributions of Risk, Journal of Exposure Analysis and Environmental Epidemiology, Volume 1, Number 4, pp 491 - 512

Clemen, 1991
Clemen, R.T., 1991, Making Hard Decisions, Duxbury Press, Wadsworth Publishing Company, Belmont, CA

Cleveland, 1993
Cleveland, W.S., 1993, Visualizing Data, AT&T Bell Laboratories, Hobart Press, Summit, NJ

Cleveland, 1994
Cleveland, W.S., 1994, The Elements of Graphing Data, AT&T Bell Laboratories, Hobart Press, Summit, NJ

Cooke, 1991
Cooke, R.M., 1991, Experts in Uncertainty, Opinion and Subjective Probability in Science, Oxford University Press, Oxford, UK

Crouch et al, 1995
Crouch, E.A.C.C., L.R. Wilson, T.L. Lash, S.R. Armstrong, and L.C. Green, 1995, Report to the Commission on Risk Assessment, Draft, Cambridge Environmental Inc., Cambridge, MA, 19 June 1995

D'Agostino & Stephens, 1986
D'Agostino, R.B. and M.A. Stephens, 1986, Goodness-of-Fit Techniques, Marcel Dekker, New York, NY

Edwards, 1992
Edwards, A.W.F., 1992, Likelihood, Expanded Edition, John Hopkins University Press, Baltimore, MD

Efron & Tibshirani, 1993
Efron, B. and R.J. Tibshirani, 1993, An Introduction to the Bootstrap, Monographs on Statistics and Applied Probability 57, Chapman & Hall, New York, NY

Evans et al, 1993
Evans, M., N. Hastings, and B. Peacock, 1993, Statistical Distributions, Second Edition, John Wiley & Sons, New York, NY

Finkel, 1990
Finkel, A.M., 1990, Confronting Uncertainty in Risk Management, A Guide for Decision-Makers, Center for Risk Management, Resources for the Future, Washington, DC, January 1990

Frey, 1992
Frey, H.C., 1992, Quantitative Analysis of Uncertainty and Variability in Environmental Policy Making, Fellowship Program for Environmental Science and Engineering, American Association for the Advancement of Science, Washington, DC

Gilbert, 1987
Gilbert, R.O., 1987, Statistical Methods for Environmental Pollution Monitoring, Van Nostrand Reinhold, New York, NY

Ibrekk & Morgan, 1983
Ibrekk, H. and M.G. Morgan, 1983, Graphical Communication of Uncertain Quantities to Nontechnical People, Risk Analysis, Volume 7, Number 4, pp 519 - 529

Israeli & Nelson, 1992
Israeli, M. and C.B. Nelson, 1992, Distributions and Expected Time of Residence for U.S. Households, Risk Analysis, Volume 12, Number 1, pp 65 - 72

Jaynes, 1982
Jaynes, E.T., 1982, On the Rationale of Maximum-Entropy Methods, Proceedings of the IEEE, Volume 70, Number 9, September 1982

Jaynes, 1957
Jaynes, E.T., 1957, Information Theory and Statistical Mechanics, Physical Review, Volume 106, Number 4, pp 620 - 630

Knuth, 1981
Knuth, D.E., 1981, The Art of Computer Programming, Seminumerical Algorithms, Volume 2, Second Edition, Addison-Wesley, Reading, MA

Kuhn, 1970
Kuhn, T.S., 1970, The Structure of Scientific Revolutions, Second Edition, University of Chicago Press, Chicago, IL

Morgan, 1984
Morgan, J.T.M., 1984, Elements of Simulation, Chapman and Hall, London, UK

Morgan & Henrion, 1990
Morgan, M.G. and M. Henrion, 1990, Uncertainty, Cambridge University Press, Cambridge, UK

NAS, 1983
National Academy of Sciences, 1983, Risk Assessment in the Federal Government: Managing the Process, National Academy Press, Washington, DC

NAS, 1991
National Academy of Sciences, 1991, Human Exposure Assessment for Airborne Pollutants, National Academy Press, Washington, DC

NAS, 1994
National Academy of Sciences, 1994, Science and Judgment in Risk Assessment, National Academy Press, Washington, DC

NCRP, 1996
National Council on Radiation Protection and Measurement, 1996, A Guide for Uncertainty Analysis in Dose and Risk Assessments Related to Environmental Contamination, NCRP Commentary, Number 14, Washington, DC

Ott, 1995
Ott, W.R., 1995, Environmental Statistics and Data Analysis, Lewis Publishers, Boca Raton, FL

Ott, 1990
Ott, W.R., 1990, A Physical Explanation of the Lognormality of Pollutant Concentrations, Journal of the Air and Waste Management Association, Volume 40, pp 1378 et seq.

Roseberry & Burmaster, 1992
Roseberry, A.M., and D.E. Burmaster, 1992, Lognormal Distributions for Water Intake by Children and Adults, Risk Analysis, Volume 12, Number 1, pp 99 - 104

Ruffle et al, 1994
Ruffle, R., D.E. Burmaster, P.D. Anderson, and H.D. Gordon, 1994, Lognormal Distributions for Fish Consumption by the General US Population, Risk Analysis, Volume 14, Number 4, pp 395 - 404

Shaw & Burmaster, 1995
Shaw, C.D. and D.E. Burmaster, 1995, Distributions of Job Tenure for US Workers in Selected Industries and Occupations, Human and Ecological Risk Assessment, in review

Shaw & Tigg, 1994
Shaw, W.T. and J. Tigg, 1994, Applied Mathematica, Addison-Wesley, Reading, MA

Smith et al, 1992
Smith, A.E., P.B. Ryan, and J.S. Evans, 1992, The Effect of Neglecting Correlations When Propagating Uncertainty and Estimating Population Distribution of Risk, Risk Analysis, Volume 12, Number 4, pp 467 - 474, December 1992

Tufte, 1990
Tufte, E.R., 1990, Envisioning Information, Graphics Press, Cheshire, CT

Tufte, 1983
Tufte, E.R., 1983, The Visual Display of Quantitative Information, Graphics Press, Cheshire, CT

Tukey, 1977
Tukey, J.W., 1977, Exploratory Data Analysis, Addison-Wesley, Reading, MA

# CASE STUDY APPLICATION: BENZENE MACT

**Michael Dusetzina**
**U.S. EPA, ORD**
**Washington, DC**

**and**

**Charles Menzie**
**Menzie-Cura Associates**
**Chelmsford, Massachusetts**

# Benzene Risk Assessment for the Petroleum Refinery MACT Standard

A Screening-Level Risk Assessment for 174 Petroleum Refineries

Risk and Exposure Assessment Group
EPA Office of Air Quality Planning and Standards

# Purpose of Assessment

- Provide more reasonable results (cancer assessment via inhalation route of exposure) than strict deterministic method
- Risk results go into a cost-benefit analysis as required for a major rulemaking (annual costs greater than $100 million)
- Results also used to determine if controls more stringent than Maximum Achievable Control Technology (MACT) are warranted; MACT level is technology driven (median control level for the best 12% of existing refineries)
- Under the 1990 CAAA, the main decision criterion regarding stringency of controls is whether exposures (and ultimately risks) to the most exposed individual are such that estimated health risks are not above a certain level

# Scope of Analysis

National analysis limited by data on health effects and refinery-specific information

- Benzene only, although 13 hazardous air pollutants (HAPs) are emitted
- 174 out of 192 identified refineries were included in the assessment (no location data on the remaining)
- Based on limited data (amount of benzene emitted by the industry was a big issue, used model plant to characterize emission releases [release height, location on plant property, etc.])
- Rulemaking under court-ordered deadline; rulemaking very high profile (regulated community involved Congress - threats of overturning the CAA)

# Brief Description of the Methodology

- Deterministic assessment completed first; risk outputs included estimated excess annual cancer incidences, number of people exposed at various risk levels, and risk to the highest exposed census block (all variables defined by an average value)
- Placed a model refinery at 172 refinery locations (latitude and longitude of each)
- Used the Human Exposure Model (version 1.5) (HEM)
  - — Industrial Source Complex (long-term) (version 2) dispersion model (yields annual average concentration estimates)
  - — U.S. Bureau of Census 1990 population data (census block basis)
  - — HEM also contains 348 meteorological stations — 5 years of meteorological data at most sites
  - — Concentrations predicted on a polar coordinate grid

# Brief Description of the Methodology (cont'd)

- Potential exposure is estimated at the center of each census block that lies between 200 and 50,000 meters of the latitude and longitude of each refinery
  - — Specifying a latitude and longitude for each refinery calls up the nearest meteorological station to the refinery location as well as the census data near the refinery
- HEM outputs fed into the Monte Carlo (Latin Hypercube) analysis; since annual cancer incidence was low, only the risks to the highest exposed census block were addressed by Monte Carlo

# Variables Used in Monte Carlo (Latin Hypercube) Assessment

- Selection of variables
  - — An important source of variability
  - — Independent of other variables in the Monte Carlo assessment
  - — Information was available to construct distribution

# Variables Used in Monte Carlo (Latin Hypercube) Assessment (cont'd)

- Variables used
  - — Residential occupancy period
  - — Breathing rate (activity level)
  - — Amount of time spent at home
  - — Amount of predicted benzene ambient concentration that enters the residential microenvironment
  - — Estimated maximum concentrations to which the residents of one census block of the thousands of potentially exposed census blocks are estimated to be exposed

# Selection of Variables and Distributions

## Residential occupancy period (years/lifetime)

CUMUL (0, 87"2, 0.1, 4, 0.25, 9, 0.5, 16, 0.75, 26, 0.9, 33, 0.95, 41, 0.98, 47, 0.99, 51, 0.995, 59, 0.999"/70

- The median value at a residence is about 9 years
- For the derivation of the benzene unit risk estimate (cancer potency) a lifetime is defined as 70 years
- In most bounding level deterministic exposure assessments for potential cancer-causing substances, 70 years of exposure at the residence is assumed (30 had been used in some assessments)
- Use of distribution for this variable significantly lowers exposure estimates
- Distribution taken from EPA's Exposure Factors Handbook

# Selection of Variables and Distributions (cont'd)

## Breathing rate (cubic meters/day)

LOGNORM (0.94, 1.44)

- A value of 20 cubic meters/day is used in the derivation of the unit risk estimate
- The most likely value from the distribution is about 18
- From TRJ Environmental analysis of Hackney data

# Selection of Variables and Distributions (cont'd)

## Amount of time spent at home (hours per week)

TRIANG (8, 16, 4, 24)24

- EPA assumption for deterministic bounding estimates had used 24 hour/day exposure assumption
- From National Human Activity Pattern Survey (EPA and others)

# Selection of Variables and Distributions (cont'd)

## Infiltration of outdoor benzene to indoor (a ratio)

TRIANG (0.72, 1, 1)

- EPA assumption for bounding assessment was that indoor concentrations were equal to outdoor
- This may not be as important as once believed
- From Mozier

# Selection of Variables and Distributions (cont'd)

## Maximum concentrations at the most exposed census block (micrograms/cubic meter annual average)

TNORMAL (38.1, 33, 7, 2, 5, 178)

- Potential concentrations from HEM output
  - — Usually 16 or 32 concentrations that bound the census block having the highest predicted impacts
  - — For example, a census block located between 200 and 500 meters from the refinery location would use the 16 concentrations estimated for 16 wind directions at 200 meters and the 16 at 500 meters for a distribution that attempts to address possible concentrations to which any person associated with the census block may be exposed

# Selection of Variables and Distributions (cont'd)

- Selected as a variable for a number of reasons
  - — Dispersion models not known for ability to predict a specific concentration at specific location/time
  - — Nearest meteorological station selected may not be representative of plant site
  - — Could potentially make very large difference in exposure estimate
- Selected truncated normal distribution
  - — A normal distribution gives negative values
  - — For Gaussian model across wind interpolation of concentrations involves an arithmetic calculation
  - — Limited curve fitting seems to be weakly lognormal

# Comment

- Perhaps use of uniform and triangular distributions are not the wisest choices; however, fairly extensive sensitivity analyses of distribution types and variable types and variable values have shown relatively little impact on the results.
- Others variables that have been described by distributions include unit risk estimates, variability in emission rates, effect of rural versus urban meteorological assumptions on predicted concentrations, variability associated with the 348 meteorological stations in database.
- Uncertainties are addressed qualitatively. By definition, zero potency is a possibility for unit risk estimates.

# Assumptions/Uncertainties/Variability

- Risk characterization focuses on benzene only, not all 13 hazardous air pollutants in petroleum.
- Effects of exposure to mixtures of compounds were not addressed.
- Sensitive subgroups were not identified; considered to be nonexistent.
- Some conservative assumptions were used, such as linearized multistage model from which unit risk estimate for benzene derived.
- Nearest airport, as opposed to onsite, meteorology was used.
- Emissions were assumed to originate at facility latitude and longitude; accuracy of these data is unknown.

# Assumptions/Uncertainties/Variability (cont'd)

- Transformation products were not addressed.
- Emissions rates were assumed to be average but not varying over time.
- Inhalation exposure only was considered.
- Uncertainty and variability not addressed separately due to resource constraints.

# BENZENE RISK ASSESSMENT FOR THE PETROLEUM REFINERY MACT STANDARD

**Michael Dusetzina**
**U.S. EPA, ORD**
**Washington, DC**

## Introduction

The purpose of this document is to highlight the results of the screening-level risk assessment for 174 petroleum refineries. This assessment was conducted as part of the regulatory impact analysis under Executive order 12886, of which cost/benefit analysis is required to support a MACT standard. The information presented here is submitted in response to a request for information on the petroleum refinery risk assessment from the National Petroleum Refiners Association.

### Organization of Document

This document contains the benzene portion of the risk assessment divided into 3 sections. The first section summarizes the hazard identification and dose-response assessment, the second section describes the exposure assessment, and the final section presents the risk characterization. Four attachments are included: Attachment 1 contains benzene emission rates (kilograms/year) for each of the 174 refineries modeled; Attachment 2 presents the HEM1.5 output for refineries that are located in ozone attainment areas; Attachment 3 displays the HEM1.5 outputs for refineries located in ozone non-attainment areas; and Attachment 4 shows modeled plant data for the Monte Carlo analysis of the refinery (one in attainment and one in non-attainment areas) associated with the highest lifetime risk (at the census block that is estimated to most at risk).

### Summary of Analytic Approach

The information presented here does not constitute a full risk characterization. Rather, it is an explanation of the methods and procedures used and a discussion of the risk associated with exposure to benzene only, as benzene risks were above the level of concern historically used by the Agency[1]. A full risk characterization would include a more thorough discussion of assumptions and will include a characterizations of all 13 emitted pollutants.

The risks from exposure to benzene presented here result from screening level assessments which use generic as opposed to site-specific parameters as inputs to the modeling. The number of facilities modeled and the impracticality (i.e. burden to industry and to taxpayers) of obtaining site-specific data for 174 facilities necessitate the use of a screening assessment.

---

[1] per the Vinyl Chloride decision (Natural Resources Defense Council v. EPA. 824 Federal Reporter F.2d 1146 (D.C. Circuit, 1987), as applied to the final NESHAP for Benzene, (55 FR 177, 1989).

The assessment includes two types of analyses: (1) a deterministic analysis based on single values, or point estimates, for the input data and (2) a probabilistic analysis that includes ranges (distributions) for certain input parameters.

The first approach (deterministic) uses single values as inputs into the exposure model and results in a point estimate of risk (e.g., $1x10^{-4}$, or 1 in 10,000 risk of cancer). The assumptions used in this analysis estimate the lifetime risk of cancer from exposure to benzene to the most exposed individual. The second analytic tool used in this analysis (probabilistic) is intended to give analysts and risk managers a better idea of the uncertainties associated with the risk estimates. This second analysis is exploratory in nature, and we recognize that it is far from complete. However, it does allow one to compare risk results based on point estimates for some of the variables with results based on distributions, and their presumed probabilities of occurrence.

The probabilistic analysis (using Monte Carlo simulation) incorporates ranges for the values of parameters that: (1) are uncertain or highly variable, (2) significantly influence the final risk estimates (i.e., those parameters to which the risk estimates are most sensitive) and (3) for which distribution data were readily available. For example, evaluated parameters include the estimated ambient concentrations, years spent in primary residence, and time spent away from the residence. This Monte Carlo analysis was conducted on a limited number of parameters, thus the analysis provides information on the variability[2]/uncertainty[3] associated with those parameters. Uncertainty and variability were not assessed separately due to resource limitations.

Both analyses focus on the risk to the individuals most exposed. This focus is appropriate for a screening-level analysis. However, this analysis is not a "worst-case" scenario. For example, several assumptions are not considered conservative such as the use of average emission rates that reflect routine (non-upset) operations. In addition, the risk estimates do not include consideration of the potential risk associated with twelve of the pollutants emitted from refineries. This risk discussion pertains only to the cancer risk from exposure to benzene.

---

[2] Variability refers to temporal, spatial, or inter-individual heterogeneity in the value of an input.

[3] Uncertainty may be thought of as a measure of the incompleteness of one's knowledge about a quantity whose true value could be determined if a perfect measuring device were available.

Summary of Results

The calculated single point values show that the leukemia risk from exposure to estimated benzene emissions from petroleum refineries ranges from a risk below one in one million to a risk above one in ten thousand. The calculated maximum individual lifetime risk exceeds one in ten thousand in both attainment and non-attainment areas.

The Monte Carlo analysis indicates that, over the range of assumptions selected, the calculated maximum risk of leukemia from exposure to benzene emissions can exceed one in ten thousand in both attainment and non-attainment areas. The calculated Monte Carlo values also include risk estimates substantially lower than the values resulting from the deterministic assessment. The calculated risk distributions for benzene are presented in the first attachment and are summarized below, along with the results of the deterministic analysis.

Summary of Maximum Individual Risk (baseline scenario)

| Location | Deterministic Value | Monte Carlo Results (range) |
|---|---|---|
| attainment | $1.6 \times 10^{-4}$ | $2.6 \times 10^{-7}$ to $7.3 \times 10^{-4}$ |
| non-attainment | $1.3 \times 10^{-4}$ | $7.3 \times 10^{-8}$ to $1.5 \times 10^{-4}$ |

Risk Distribution (Deterministic; assuming exposure to benzene only)

| Baseline Risk Level | People at or above risk level | Number of refineries at or above risk level: National | Number of refineries at or above risk level: Ozone Attainment Areas |
|---|---|---|---|
| 1x10-4 | 514 | 7 | 3 |
| 1x10-5 | 89,900 | 84 | 43 |
| 1x10-6 | 4,481,000 | 153 | 79 |
| Floor | | | |
| $1 \times 10^{-4}$ | 152 | 3 | 1 |
| $1 \times 10^{-5}$ | 61,400 | 62 | 27 |
| $1 \times 10^{-6}$ | 3,068,000 | 108 | 73 |
| Proposed Rule | | | |
| $1 \times 10^{-4}$ | 19 | 2 | 1 |
| $1 \times 10^{-5}$ | 40,400 | 57 | 25 |
| $1 \times 10^{-6}$ | 1,880,000 | 101 | 70 |

Since these values are estimates of the risk associated only with leukemia formation from the exposure to benzene emissions, we believe that the maximum risk estimate may be an underestimate of the total risk from exposure to emissions from petroleum

refineries. The total risk would also include consideration of non-quantifiable risks from exposure to the other pollutants emitted from refineries as well as other non-quantifiable risks associated with benzene exposure.

$$\text{Risk} = \underset{\text{quantified}}{\text{benzene risk}} + \underset{\text{non-quantified}}{\text{benzene risk}} + \underset{\text{non-quantified}}{\text{other pollutant risk}}$$

We consider this analysis an appropriate application of Monte Carlo simulation, and believe that it provides decision-makers with important information beyond that available from point estimates alone. We understand that using different distributions or evaluating alternate combinations of parameters may result in different final results. Our goal here is to conduct what we consider to be a sound analysis that can ultimately be used by EPA decision-makers to make appropriate risk management decisions. We invite thoughtful comment on this analysis and encourage dialogue that would result in improvements to the methods.

## Hazard Identification and Dose Response Assessment Summary

Thirteen species of HAP's were identified as emissions from the evaluated facilities. The benzene emissions are of the greatest concern in terms of cancer risk; thus this discussion focuses solely on the potential for increased risk of cancer due to exposure to benzene emissions from the facilities evaluated.

Benzene is classified as a known human carcinogen. There is sufficient human epidemiological evidence to support the claim that exposure causes an increased risk of cancer to humans. Benzene is of particular concern to EPA because long-term exposure to this chemical has been shown to cause leukemia in humans. While this is the best known effect, benzene exposure is also associated with aplastic anemia, multiple myeloma, lymphomas, pancytopenia, chromosomal breakages, and weakening of bone marrow. A reduction in human exposure to benzene could lead to a decrease in cancer risk and ultimately to a decrease in cancer mortality.

The quantitative dose response information the Agency uses to address benzene is in the form of a unit risk estimate (URE). A URE usually represents a plausible upper bound of the increased risk of developing cancer for an individual continuously exposed throughout a lifetime (70 years) to one unit (defined as 1 microgram per cubic meter (ug/cu.m)) of the potential carcinogen in the air. Some UREs are based on animal studies which are extrapolated to humans; others are based on human data. The URE for benzene was based on human data from an occupational setting. The Agency has higher confidence in UREs based on human data.

Risks calculated using an upper bound URE are not expected to be any higher than the predicted numbers and may be substantially lower, including a risk of zero.

## Exposure Assessment Methodology

This section explains the methodology used to estimate individual and population exposure from inhalation of benzene emitted from petroleum refineries.

A screening risk assessment was conducted for 174 of 192 petroleum refineries. The assessment was conducted as part of the regulatory impact analysis under Executive order 12886, which includes cost/benefit analysis, to support the MACT standard. The risk characterization is a screening level assessment because the analysis used generic, as opposed to site-specific, parameters as inputs to the modeling. For example, information was not available on local meteorological conditions at the refineries so meteorological data from the nearest meteorological station (airport) were used to represent average conditions at the refinery. Similarly, because the precise location of specific emission sources (releases) on plant property were not available, the emissions were assumed to originate from the center of the location provided for each refinery (latitude and longitude). Emissions from eighteen refineries were not modeled since location information either was judged to be inaccurate or was not available.

For the analysis, the refineries were divided into two groups: those located in ozone non-attainment areas and those located in attainment areas. Two types of analyses were also used in the assessment to address both groups. The first consisted of a bounding analysis where all input variables were described by individual values (point estimates). This is a deterministic procedure that produces outputs of single numbers rather than ranges. The second approach incorporates into the analysis ranges of values for those variables meeting the following criteria: mathematical distributions are available; the variables are independent; and, most importantly, the variables are believed to significantly influence the results of the analysis. This probabilistic procedure uses Monte Carlo simulation to produce distributions with associated probability estimations (e.g. there is a 95% probability that the estimated risk to the most exposed population group (census block) is less than one in ten thousand).

The distributions used in the Monte Carlo analysis were taken primarily from EPA sources (such as the Exposure Factors Handbook) and the literature. Best judgments were used in selecting the distributions and, in some cases, in using only portions of the distributions that are provided in the Handbook. Use of other distributions may result in different final outcomes for the Monte Carlo analysis.

The model used for the deterministic, screening assessment is referred to as the Human Exposure Model (HEM version 1.5). The

model was run separately for refineries in ozone non-attainment areas and those in attainment areas. Three risk measures are produced from the model. The first is a risk distribution (i.e., numbers of people at or above a specified risk level). The second is an estimate of the annual cancer incidence, or the expected additional cancer incidence (leukemias) per year, from exposure to the petroleum refinery emissions. Last is an estimate of the risk to the census block of people exposed to the highest benzene concentrations predicted near their residence (called maximum individual risk (MIR)).

The model assumes that exposure occurs at the center of a U.S. Bureau of Census population block. Even if a census block abuts a refinery, exposure is not assumed to occur at the refinery fenceline. This approach will likely underestimate the exposures experienced by those residents who live closer to the refinery than the location of the population center. The extent of the underestimation depends on the size of the affected block. The underestimate will tend to be in rural area because blocks are generally smaller in size in urban areas (but have larger populations) than in rural areas. Attachment 1 presents the results of the analyses.

Because meteorological data collected on site by the refinery are usually not available, the dispersion model that is included in HEM 1.5 uses the meteorological data recorded at the weather station closest to each refinery. In other words, if a refinery is located in Pittsburgh, then the meteorological data from the Greater Pittsburgh International Airport would be used for this refinery. The airport data represents the average of five years of data.

Since data on the location of emission sources on plant property were not available, emissions were modeled as if they originated from the location (latitude and longitude) that represents each refinery. Emissions were assumed to be released from a height of 3 meters, at an exit velocity of 0.01 meters per second, and at an exit temperature of 295 degrees Kelvin. These generic assumptions about the emission releases are believed to represent the typical conditions for the refinery sources that were modeled. A sensitivity analysis of the effect of representing emissions as a point source rather than an area source was conducted. This analysis indicates that the predicted concentrations associated with the MIR would be approximately ten percent higher if refinery emissions were modeled as an area source covering a 100 meter square area rather than modeled as a point source. Attachment 1 presents the benzene emission rates for each refinery.

Note that the data used as inputs into the risk assessment contained speciated HAP (hazardous air pollutant) emissions

calculated from the equipment leak emissions for each of the 192 refineries in the U.S. Because speciated HAP data was not available for the three other types of emissions (tanks, waste water, and process vents), benzene emissions per refinery were calculated by dividing the equipment leak emissions with the fraction of total refinery HAP emissions attributed to equipment leaks. The emission estimates for equipment leaks will be revised based on newly submitted data, and questions related to process emissions and storage tanks are being addressed. It is not certain what effect the emission rate revisions will have; however, a significant change in benzene emission estimates is not expected, and thus no significant change in the estimated risks is expected.

## Risk Characterization

The results shown below present the MIR for both the deterministic and Monte Carlo approaches. Estimates of annual cancer incidence are based entirely on the deterministic approach. A deterministic (point estimate) and a probabilistic approach (Monte Carlo) were used for MIR for refineries located in both attainment and non-attainment areas. Annual cancer incidence was calculated separately for all refineries that are located in attainment and all located in non-attainment areas. The MIR, the highest individual risk, usually results from a refinery with relatively large benzene emissions and with a population block near the emissions source. Attachments 2 and 3 present results of the deterministic analyses for attainment and non-attainment areas, respectively.

Summary of Results

(The number of significant figures presented are only for comparing different values, not for demonstrating accuracy.)

Maximum individual risk (baseline, benzene) using EPA's unit risk estimate

| Location | Deterministic Value | Monte Carlo Results (range) |
|---|---|---|
| Attainment | $1.6 \times 10^{-4}$ | $2.6 \times 10^{-7}$ to $7.3 \times 10^{-4}$ |
| Non-attainment | $1.3 \times 10^{-4}$ | $7.3 \times 10^{-8}$ to $1.5 \times 10^{-4}$ |

Annual cancer incidence (baseline, benzene)

| | |
|---|---|
| Attainment | 0.064 |
| Non-attainment | 0.27 |

The risk characterization (inhalation route of exposure only) was conducted for 174 of 192 refineries that are located in many states. Thirteen listed HAPs have been identified in the emissions from these facilities; however, this risk characterization focuses only on benzene. Benzene is classified as a class A carcinogen (based on human data). The URE or cancer potency estimate is $8.3 \times 10^{-6}$. About 84 million people live within 50 kilometers of these refineries and are therefore potentially exposed to emissions from these facilities.

The estimated deterministic (single number) results for potential carcinogenic effects for benzene show an MIR value of 1.6 chances out of 10,000. In other words, the people that reside at the census block that experiences the highest risks from benzene equipment leak emissions from any refinery could have 1.6 chances out of 10,000 of developing leukemia. The estimated annual incidence from benzene equipment leak emissions from the 174 refineries modeled was less than one (specifically, the annual incidence is estimated to be 0.3).

The following table contains the estimated (deterministic) number of people and refineries (nationally, and for those located in attainment areas) that are exposed to cancer risks that equal or exceed various risk levels. The table also shows the effect of further controlling equipment leak benzene emissions.

Risk Distribution (Deterministic, assumes exposure to a single pollutant; uses EPA's unit risk estimate)

| | Health Risk - Benzene | | |
|---|---|---|---|
| | | Number of refineries at or above risk level | |
| Baseline Risk Level | People at or above risk level | National | Ozone Attainment Areas |
| 1x10-4 | 514 | 7 | 3 |
| 1x10-5 | 89,900 | 84 | 43 |
| 1x10-6 | 4,481,000 | 153 | 79 |
| Floor | | | |
| $1x10^{-4}$ | 152 | 3 | 1 |
| $1x10^{-5}$ | 61,400 | 62 | 27 |
| $1x10^{-6}$ | 3,068,000 | 108 | 73 |
| Proposed Rule | | | |
| $1x10^{-4}$ | 19 | 2 | 1 |
| $1x10^{-5}$ | 40,400 | 57 | 25 |
| $1x10^{-6}$ | 1,880,000 | 101 | 70 |

Exploratory Monte Carlo Analysis

A Monte Carlo approach was used to characterize the distribution of possible results given the uncertainty in several key assumptions. The Monte Carlo analyses for MIR included two facilities; the one facility that is in an ozone attainment area and the one in a non-attainment area that cause the highest MIR compared to the other refineries that were modeled.

Estimates of breathing rates, and years spent in the primary residence were varied according to estimates from the American Industrial Health Council's (AIHC) Exposure Factors Sourcebook[2]. Estimates of time spent away from home were taken from EPA's Exposure Factors Handbook[3]. Estimates of the fraction of the predicted outdoor benzene concentrations that infiltrate to indoor microenvironments were based on two estimates; one from the literature[4], and one based on scientific judgement.

The Monte Carlo analysis takes as input the distribution of key assumptions and produces as a result a probabilistic estimation of the range of risk estimates, given the initial data characterization. In this way EPA is able to determine the probability associated with any value on the distribution including the median estimate of risk and its plausible highest value.

The input assumptions evaluated using this technique are ambient concentrations, breathing rates, time spent away from home, years spent at the primary residence, and building infiltration rate. The assumed distributions for these variables follow.

The distributions shown and the format used is for conducting Monte Carlo analyses using personal computer software by @RISK from Palisade, Corp. (607-277-8000).

Breathing rates (cubic meters per day)

TRIANG(6,18.9,32)/20 from Exposure Factors Sourcebook[2].

The distribution is assumed to be triangular, with a minimum value of 6 cubic meters per day, a most likely value of 18.9, and a maximum value near 32. The end values of a triangular distribution have a zero probability of being sampled. Including this distribution allows one to consider the effect of the variability in breathing rates across the population as well as variability within an individual. Each value sampled is divided by 20 to give the result. The breathing rate assumed in deriving the benzene URE is 20.

Time spent away from home (hours per week)

CUMUL(0,107,"0.34,0.3,8.3,0.4,20.2,0.5,32.1,0.6,37.7,0.7,41.3,0.8,46.9,0.9")/168

from EPA's Exposure Factors Handbook[3]

This is a cumulative distribution with a minimum value of zero hours per week, a maximum value of 107 hours per week, an initial value of 0.34 hours with a frequency of 30%, 8.3 hours at

40%, 20.2 hours at 50%, 32.1 hours at 60%, 37.7 hours at 70%, 41.3 hours at 80%, and 46.9 hours at 90%. There are 168 hours in a week and each sampled value is divided by 168 to give the result. This distribution recognizes that most people do not remain in one location during any one day and therefore are not subject to the same exposure for a continuous period. Exposure is assumed to be zero when away from home.

Residential occupancy (years spent in primary residence)

CUMUL(1,75,"4,0.25,9,0.5,16,0.75,26,0.9,33,0.95,47,0.99)/70

from Exposure Factors Sourcebook[2]

This is a cumulative distribution with a minimum value of one year, a maximum value of 75 years, 25% of homeowners live in their residence 4 years or less, for 9 years the value is 50%, for sixteen years - 75%, for 26 years - 90%, for 33 years - 95%, for 47 years - 99%. The URE is based on lifetime exposure. For this analysis a lifetime is assumed to be 70 years and thus each sampled value is divided by 70 to give the result. Exposure is assumed to be zero after change of residence.

In/Out (a ratio) infiltration of outdoor concentrations to indoor microenvironments

UNIFORM(0.4,1.0) from literature[4], scientific judgement

This is a uniform distribution with a minimum value of 0.4 and a maximum value of one. All values within this range have the same probability of being selected during sampling. This parameter accounts for the variation in exposure the exposed population faces in various microenvironments such as school, work, outdoors, etc. Due to differing air exchange rates, weather patterns, the indoor concentration may reasonably be considered to be higher than 0.4 as reported in the literature. Here, we used scientific judgement to set the upper end of the range at 1, which assumes that outdoor and indoor concentrations of benzene are equal, as indoor concentrations could equal outdoor concentrations where air exchange rates are high (for example in some air conditioned buildings).

Estimated ambient concentrations to which the MIR group is potentially exposed ($ug/m^3$ annual average)

Attainment area

TNORMAL(35.2,35.1,3.0,146) from HEM output, location of most exposed census population block; see attachment 5 refinery 138

This is a truncated normal distribution with a mean of 35.2 micrograms per cubic meter annual average concentration, a standard deviation of 35.1, ranging between 3.0 and 146 micrograms per cubic meter. This variable was selected because the location of the population associated with the census block receiving the highest exposures is uncertain. The only information known is the location of the area-weighted center of this population block and the number of people assigned to the block. This distribution is intended to capture possible concentrations to which people residing in the most exposed census block are exposed. Attachment 5 shows the location to be between 50 meters and 100 meters from the latitude and longitude locating the benzene emission sources. The rule of thumb is to use those rings of estimated ambient concentrations that bound the centroid (i.e., 50 meters and 100 meters) in this distribution. However, most refineries are rather large in area, concentrations at the 200 meter ring were also used because the accuracy of the refinery's latitude and longitude is not known. The effect was to lower the mean but still allow consideration of all possible concentrations that the exposed MIR population may be exposed to.

Non-attainment area

| | |
|---|---|
| TNORMAL(13.4,13.1,1.53,51.2) | from HEM1.5 output and location of most exposed census block; see attachment 5 refinery 172 |

Attachment 4 shows the most exposed census block to fall between 100 and 200 meters from this refinery. Thus, concentrations at 100 and 200 meters were used to define this distribution.

Results:

The number of significant figures presented are only for comparing different values, not for demonstrating accuracy.

Maximum Individual Risk for refiners located in ozone attainment areas using EPA's unit risk estimate

| Cumulative Probability | MIR |
|---|---|
| minimum to maximum | $2.6 \times 10^{-7}$ to $7.3 \times 10^{-4}$ |
| 5% to 95% | $2.1 \times 10^{-6}$ to $1.4 \times 10^{-4}$ |
| 50% (median) | $2.1 \times 10^{-5}$ |
| 90% | $9.8 \times 10^{-5}$ |
| 95% | $1.4 \times 10^{-4}$ |
| 98% | $2.2 \times 10^{-4}$ |

The unit risk estimate for benzene is considered to be one of the parameters in the analysis with the most associated uncertainty; however, data are not available to develop a probability distribution for the benzene unit risk estimate. To provide an indication of the range of maximum risk estimates that one could see if other plausible unit risk estimates were used, we conducted additional analyses for two alternative benzene unit risk estimates. One unit risk estimate was developed by the American Petroleum Institute, using a quadratic dose-response model rather than the linearized multi-stage model that the EPA uses. The second alternative unit risk estimate was developed by the California Air Resources Board and is characterized as an upper 95th percentile estimate based on animal data. The EPA unit risk estimate is developed from a maximum-likelihood estimate. It represents the central tendency estimate from the dose-response curve based on human data; however, because the dose-response model is based on several conservative assumptions (e.g., linearity at low doses), the EPA's unit risk estimate is characterized as an upper bound estimate.

Using the American Petroleum Institute URE[5] ($3.2 \times 10^{-7}$) the MIR ranged from $1.0 \times 10^{-8}$ to $2.8 \times 10^{-5}$. The California Air Resources Board URE[6] ($2.9 \times 10^{-5}$) gives a range of $9.1 \times 10^{-7}$ to $2.6 \times 10^{-3}$.

Maximum Individual Risk for refiners in ozone NON-attainment areas using EPA's unit risk estimate

| Cumulative Probability | MIR |
|---|---|
| minimum to maximum | $7.3 \times 10^{-8}$ to $1.5 \times 10^{-4}$ |
| 5% to 95% | $7.7 \times 10^{-7}$ to $5.3 \times 10^{-5}$ |
| 50% (median) | $8.1 \times 10^{-6}$ |
| 90% | $3.9 \times 10^{-5}$ |
| 95% | $5.3 \times 10^{-5}$ |
| 98% | $1.5 \times 10^{-4}$ |

Using API's URE the range is estimated to be $2.8 \times 10^{-9}$ to $5.8 \times 10^{-6}$, and using CARB's estimate the range is $2.5 \times 10^{-7}$ to $5.3 \times 10^{-4}$.

The use of MC and values that are commonly used in similar MC assessments have enabled the Agency to conduct the refinery risk characterization in a way that is perceived by many to be less conservative than the traditional "point estimate" approach. As discussed previously, for a number of reasons, quantitative analysis of the total risk from exposure to emitted pollutants other than benzene was not conducted. The risks from exposure to benzene alone, exceed 1 in 10,000 for the maximum exposure scenario.

Discussion of Assumptions/Uncertainties/Variability

The degree to which any risk estimate represents the true risk depends on a number of factors including the assumptions used in the analysis, and the uncertainty and variability present. These factors may either under- or over-estimate the true risk. A brief discussion of some of these factors follows.

In this analysis, total risks may be higher than predicted because chemicals are emitted for which we have no health benchmarks and were left out of the analysis (i.e., risks are assumed to be zero for those compounds without health benchmarks). Sensitive population groups may exist, but could not be identified, and thus were considered to be non-existent. Effects of being exposed to mixtures of compounds are not addressed.

Risks may be lower than estimated due to the use of some conservative assumptions such as the linearized multistage model from which the URE for benzene is derived.

Other assumptions can either increase or decrease the estimated risk results. These include, but are not limited to, the use of nearest airport versus on-site meteorology and assuming emissions originate from the latitude and longitude that locates the refinery versus where they actually occur on plant property. Transformation products (which could increase or decease risk) were not addressed. Also, this analysis assumed that the emission rates provided by ESD were average but did not vary over time.

This analysis considered inhalation exposure only; however, non-inhalation exposure risks for benzene are not believed to be significant. The analysis addressed only part of the overall uncertainty that is inherent in risk assessments (e.g. emission rates). Uncertainty and variability were not addressed separately in this analysis due to resource constraints.

In general, the accuracy of the facility locations (latitudes and longitudes) of the refineries are unknown. For other source categories location data has been historically weak. Inaccurate location data can significantly impact the MIR estimation and to a generally lessor extent can effect annual incidence.

Conclusion

The calculated single point values show that the leukemia risk from exposure to benzene emissions from specific sources at petroleum refineries ranges from a risk below one in one million to a risk above one in ten thousand. The calculated maximum individual lifetime risk exceeds one in ten thousand in both attainment and non-attainment areas.

The Monte Carlo analysis indicates that, over the range of assumptions selected, the calculated maximum risk of leukemia from exposure to benzene emissions can exceed one in ten thousand in both attainment and non-attainment areas. The calculated Monte Carlo values also include risk estimates substantially lower than the values resulting from the deterministic assessment.

References


1. Frey, C. "Distribution Development for Probabilistic Exposure", Paper No. 95-TA42.02, Air and Waste Management Association Annual Meeting, San Antonio, Texas, June 1995.

2. Exposure Factors Sourcebook. American Industrial Health Council. May 1994.

3. Exposure Factors Handbook. Office of Health and Environmental Assessment. U.S. Environmental Protection Agency. EPA/600/8-89/043. July 1989.

4. Johnson, T., et al, Estimation of Incremental Benzene Exposures and Associated Cancer Risks Attributable to a Petroleum Refinery Waste Stream Using the Hazardous Air Pollutant Exposure Model (HAPEM), Paper No. A1389, 86th Annual Meeting of the Air and Waste Management Association, Denver, CO. June 1993.

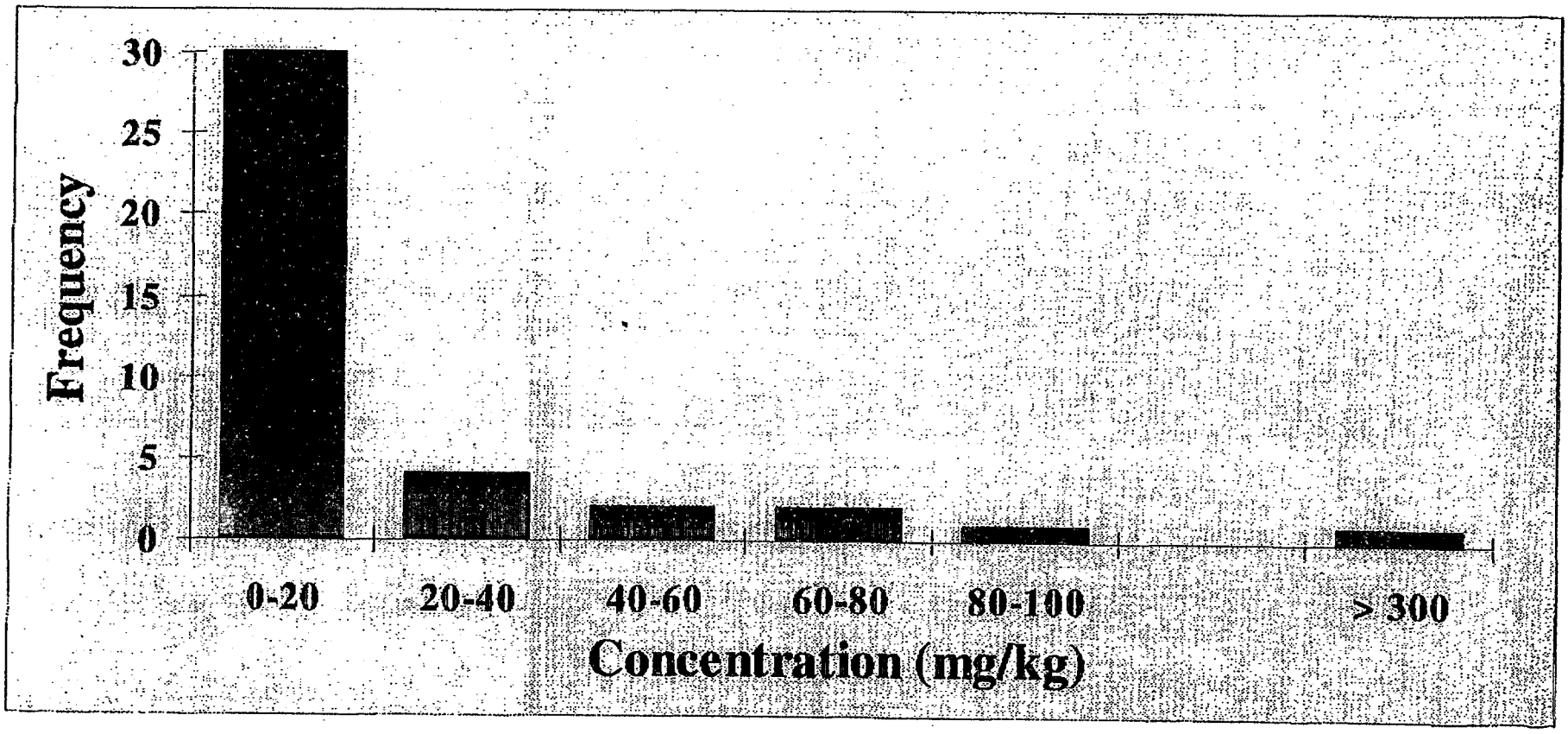
5. Yosie, T.F., American Petroleum Institute submittal on the benzene-induced risk of leukemia; letter with attachment. U.S Environmental Protection Agency Docket A-79-16. October 1988.

6. Air Toxics "Hot Spots" Program Risk Assessment Guidelines. California Air Pollution Control Officers Association (CAPCOA). January 1992.

# CASE STUDY APPLICATIONS SUPERFUND SITE

**Teresa Bowers**
**Gradient Corporation**
**Cambridge, Massachusetts**

# PCB Sampling Data



# Summary Statistics

N = 40

$\overline{x}$ = 23.8

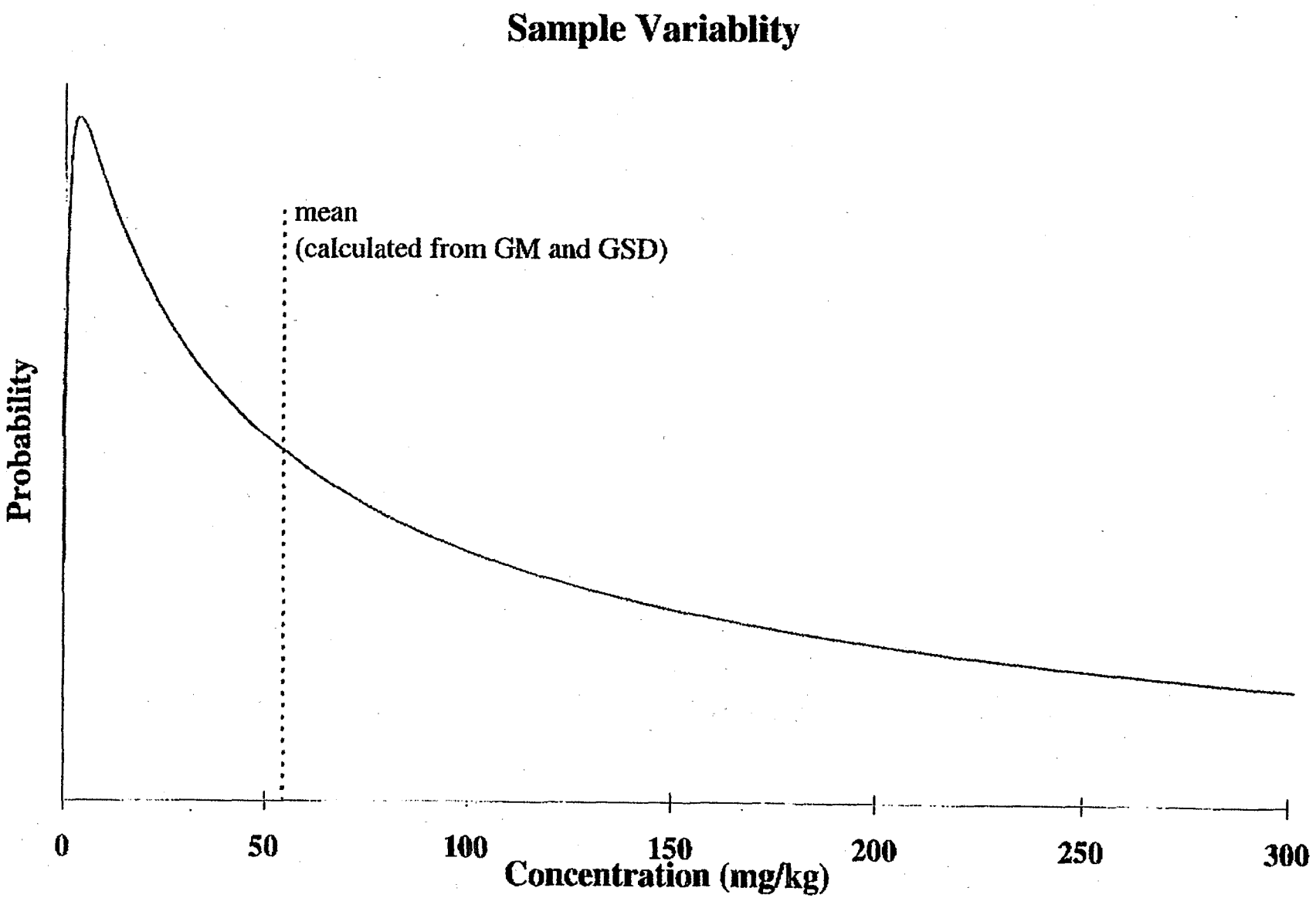
s = 58.9

GM = 3.7

GSD = 10.2

max = 360

95% UCL on mean = 252

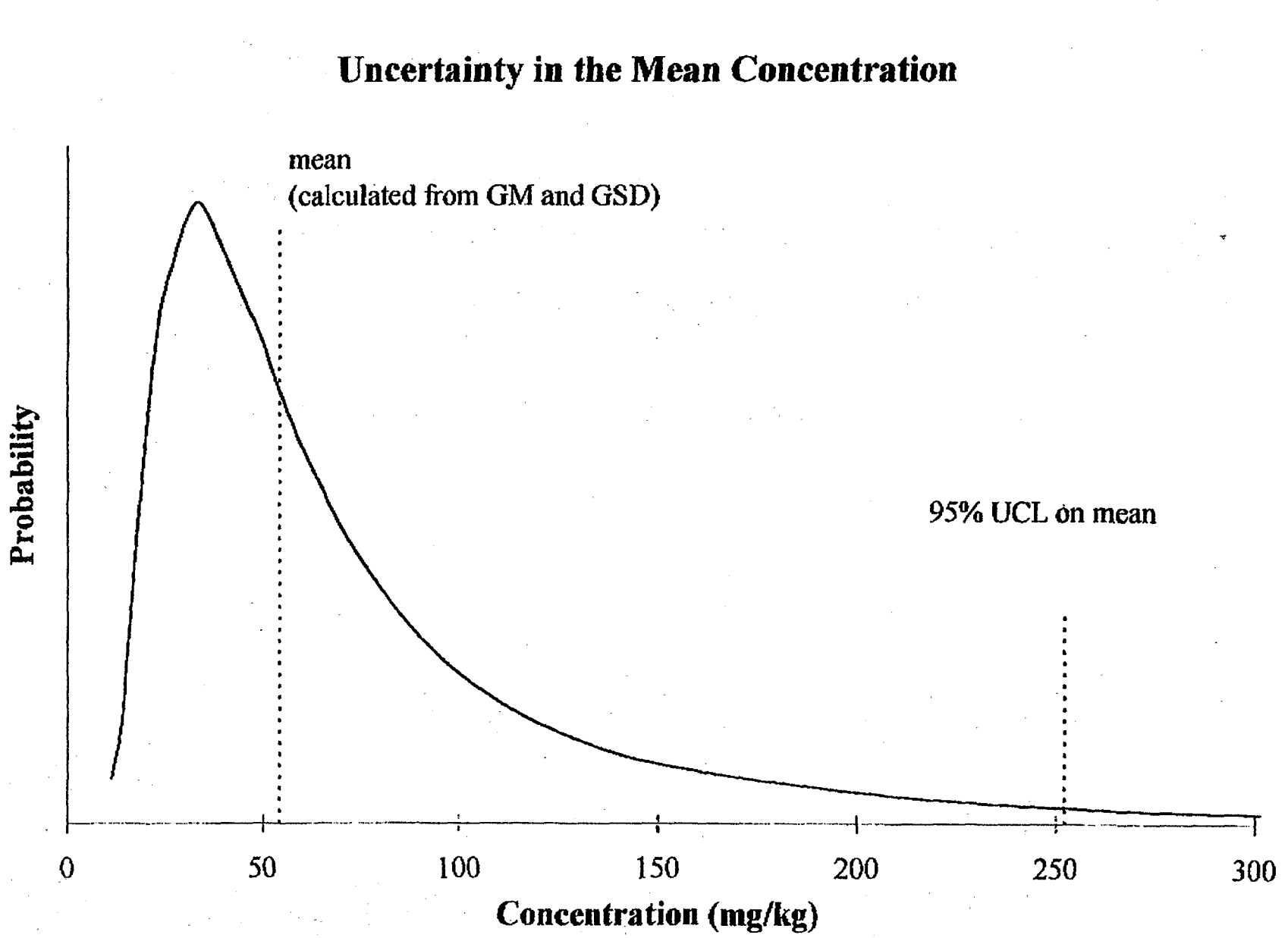

Sample Variablity
mean
(calculated from GM and GSD)
Probability
0
50
100
150
200
250
300
Concentration (mg/kg)

# Calculation of Confidence Limits on Mean Concentration

$$CL_{\alpha} = e^{\mu + \frac{\sigma^2}{2} + \frac{\sigma H}{\sqrt{N-1}}}$$

**Where:**

- $CL_{\alpha}$ is the confidence limit on the arithmetic mean at confidence level $\alpha$
- $\mu$ = log (geometric mean), calculated from the concentration data
- $\sigma$ = log (geometric standard deviation), calculated from the concentration data
- H = H-statistic
- N = number of samples


Uncertainty in the Mean Concentration
mean
(calculated from GM and GSD)
95% UCL on mean
Probability
0
50
100
150
200
250
300
Concentration (mg/kg)

Sample Variability
mean
(calculated from GM and GSD)
Probability
0
50
100
150
200
250
300
Concentration (mg/kg)

# QUANTITATIVE TECHNIQUES FOR ANALYSIS OF VARIABILITY AND UNCERTAINTY IN EXPOSURE AND RISK ASSESSMENT

**Christopher Frey**
**North Carolina State University**
**Raleigh, North Carolina**

# *Quantitative Techniques for Analysis of Variability and Uncertainty in Exposure and Risk Assessment*

H. Christopher Frey, Ph.D.
Assistant Professor
Department of Civil Engineering
North Carolina State University
Raleigh, NC 27695-7908

Workshop on Monte Carlo Analysis
U.S. Environmental Protection Agency
New York City

April 15, 1996

## OUTLINE

- Distinguishing Between Variability (V) and Uncertainty (U)
- Developing Distributions
- Dependences Among V & U
- Simulation
- Analysis of Results
- Presentation of Results

NC STATE

## VARIABILITY AND UNCERTAINTY

- Uncertainty
  - lack of knowledge
  - Stochastic variability, "Type B Uncertainty"
  - probability distribution
- Variability
  - heterogeneity in time, space, etc.
  - "Type A Uncertainty"
  - frequency distribution
- Can have certainty about variability

NC STATE

## VARIABILITY VS. UNCERTAINTY

- Uncertainty: How probable is it that a risk will be over- or under-estimated?
- Variability: Certainty that different individuals will be subject to different risks.
- Can "reduce" variability by disaggregation
  - Stratify into more homogenous groups
    - » Ex: pica children
  - Identify sensitive subpopulations
- Can reduce uncertainty by additional research
  - Collect more data
  - Obtain better measurements
  - Trade-off with cost

NC STATE

## INDIVIDUAL VS. POPULATION RISK

- Variability and Uncertainty co-mingled: Risk to an individual selected at random from the population
  - Not very useful in most regulatory contexts
- Uncertainty in the risk to a highly exposed individual
  - Uncertainty regarding a fractile of the population
- Uncertainty in the population risk
  - Uncertainty regarding the average for the population
- For most regulatory purposes, a distinction between variability and uncertainty is important

NC STATE

## UNCERTAINTY ABOUT VARIABILITY

- Measurement Error
- Small sample sizes
- Non-representative samples
- Model error (e.g., in selecting a parametric distribution)

NC STATE

## Interactions Between V & U

$$X_{i,j} = X_{v,i}(\hat{\theta}_{v,j}) + X_{u,j}(\hat{\theta}_{u,i})$$

$X_{v,i}(\hat{\theta}_{v,j})$ = Family of frequency distributions

$\hat{\theta}_{v,j}$ = Sampling distributions for parameters

$X_{u,j}(\hat{\theta}_{u,i})$ = Probability distributions for measurement error

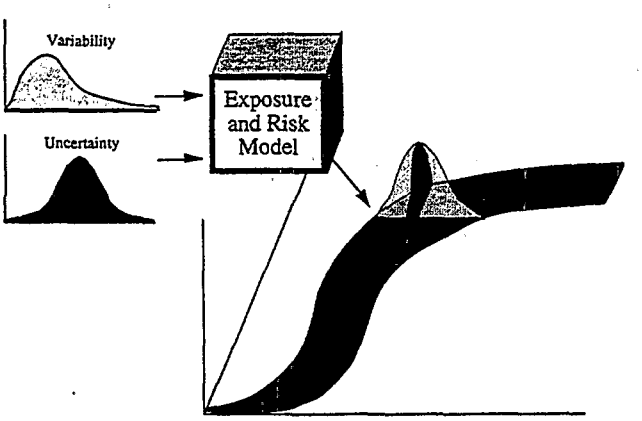
$\hat{\theta}_{u,i}$ = Parameters of probability distributions

i = uncertainty index (1,n)

j = variability index (1,m)

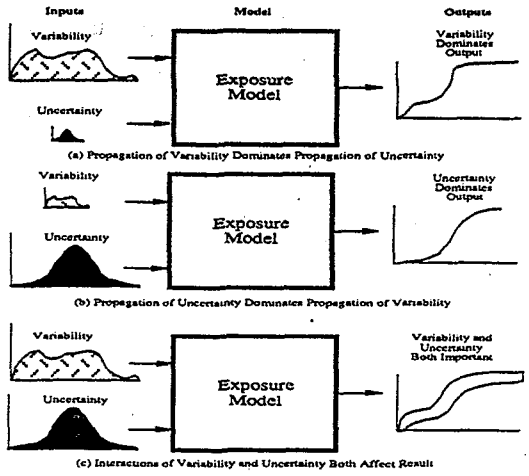
NC STATE

## VARIABILITY AND UNCERTAINTY




NC STATE

# INTERACTIONS BETWEEN VARIABILITY AND UNCERTAINTY



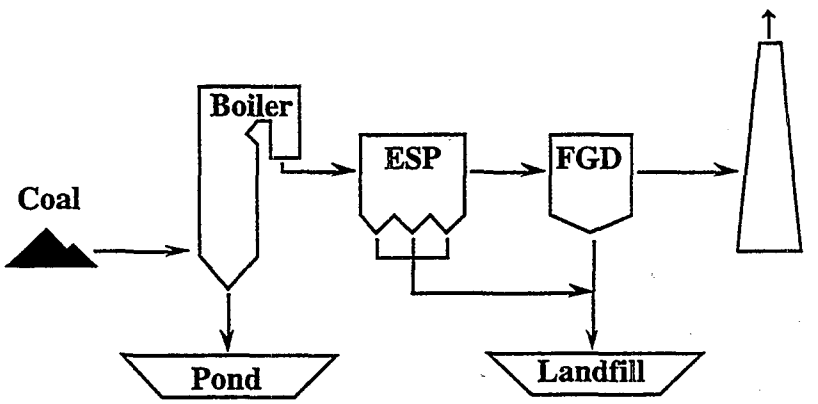

NC STATE

# Distribution Development

- Uncertainty due to small sample size
- Example:
  - Partitioning Factor
  - 5 data points

NC STATE

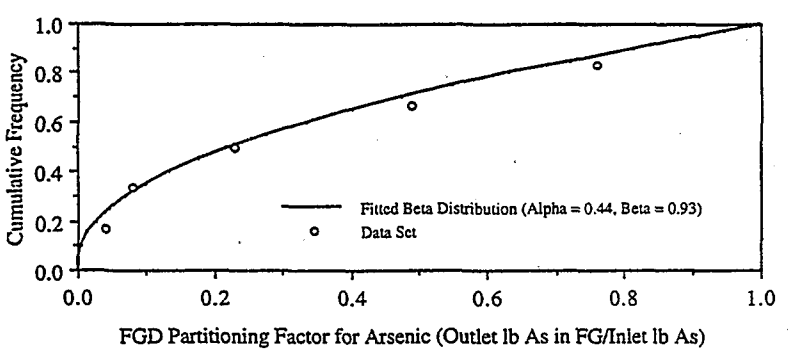
## Illustrative Case Study: Arsenic Emissions from a Power Plant




NC STATE

## Typical Approach

- Fit a Frequency Distribution to Data




NC STATE

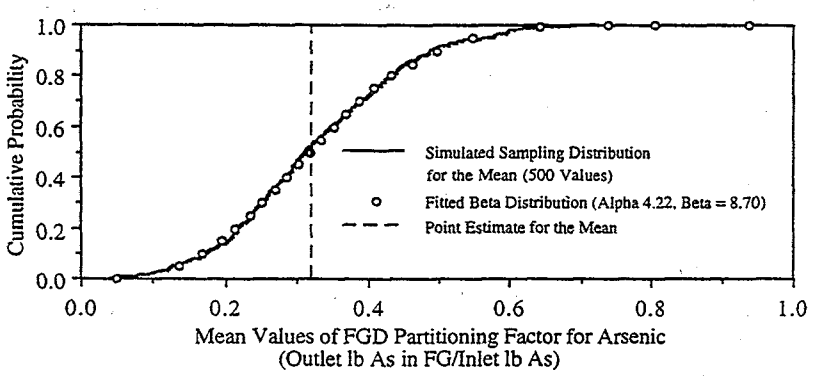
## Uncertainty due to Sampling Error

Simulating Sampling Distributions and Their Dependences

- Simulate sets of samples drawn from parent frequency distribution
- Calculate mean, variance, or other parameters for each sample
- Repeat
- Develop simulated sampling distributions for parameters

NC STATE

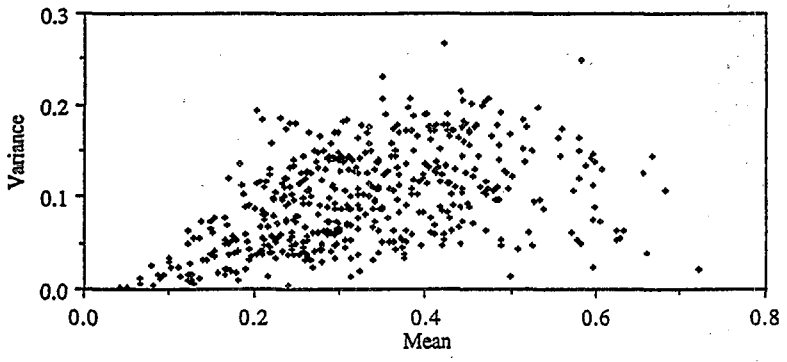
## Example: Uncertainty in Mean




*Large Range of Uncertainty in Mean Due to Having Only Five Data Points*

NC STATE

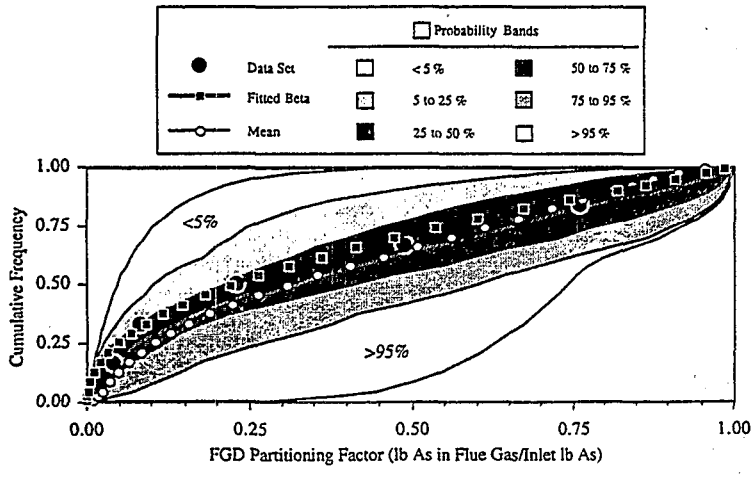
## Dependence Between Sampling Distributions




*Assumptions of Independence or Linear Dependence Would Not be Valid*

NC STATE

## Interactions Between Variability and Uncertainty (Sampling Error)




NC STATE

## Uncertainty Due to Measurement Error

- Measurements Comprised of:
  - Frequency distribution for variability
  - Probability distribution for random error
  - Bias due to systematic error
- $M = f_{c,i} + E$, where $E \sim N(\mu_E,$
- $\mu_{fc,i} = \mu_M - \mu_E$
- $(\sigma_{fc,i})^2 = (\sigma_M)^2 - (\sigma_E)^2$

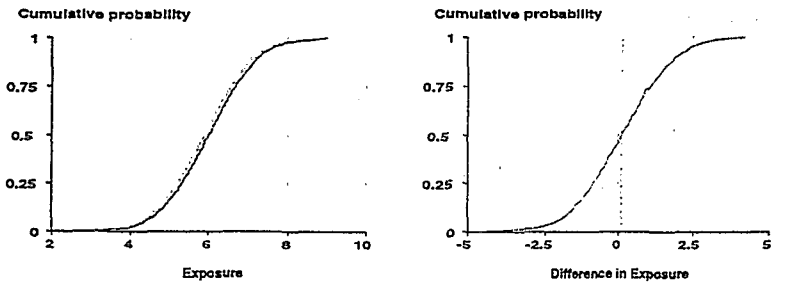
NC STATE

## Dependences

- Between Variable Quantities
  - e.g., Intake Rate, Body Weight
  - Mechanistic or empirical models preferred
- Between Uncertain Quantities
  - Sampling Distributions
  - Measurement Errors
- Between Uncertain and Variable Quantities
  - Measurement error = f(variable quantity)
  - Frequency distribution = f(sampling distributions)

NC STATE

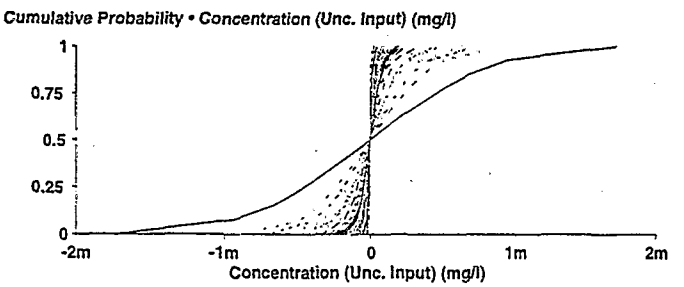
## Dependent vs. Independent Measurement Error




- Two individuals with slightly different exposures
- If $\rho$=1, one individual dominates the other
- If $\rho$=0, 48% probability individual with lower mean has a higher exposure (changes rank order)

NC STATE

## UNCERTAINTY AS A FUNCTION OF VARIABILITY




Standard deviation of concentration uncertainty depends on magnitude of concentration

NC STATE

## MODELING

- Analytical Methods
- First Order Methods
- Approximation Techniques
- Numerical Methods

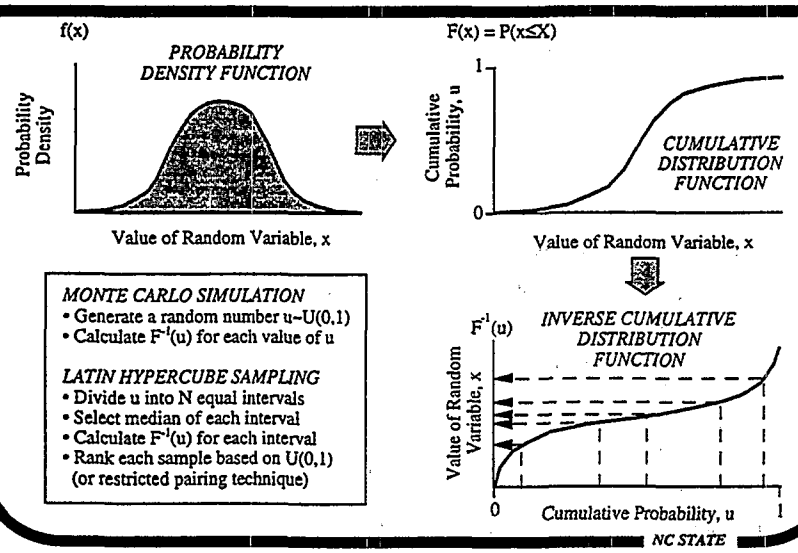
NC STATE

## MODELING

- Numerical Methods
  - Monte Carlo simulation
  - Latin Hypercube Sampling (LHS)
  - Other sampling techniques

NC STATE

# MONTE CARLO SIMULATION

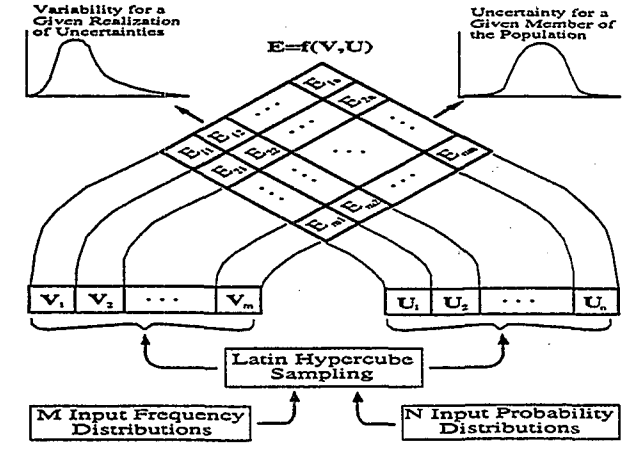



*MONTE CARLO SIMULATION*
- Generate a random number u~U(0,1)
- Calculate $F^{-1}(u)$ for each value of u

*LATIN HYPERCUBE SAMPLING*
- Divide u into N equal intervals
- Select median of each interval
- Calculate $F^{-1}(u)$ for each interval
- Rank each sample based on U(0,1) (or restricted pairing technique)

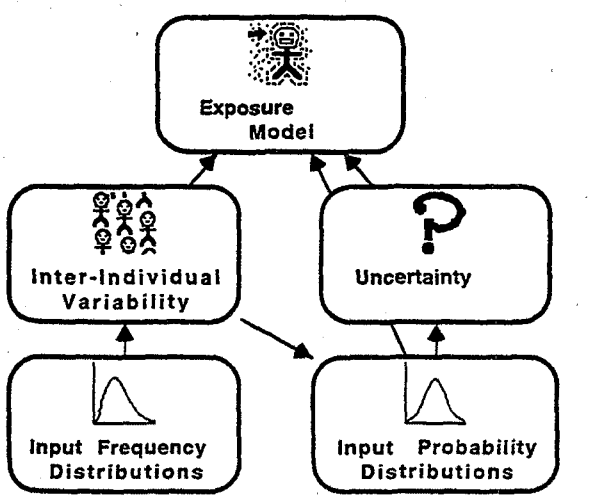
NC STATE

# TWO-DIMENSIONAL SIMULATION



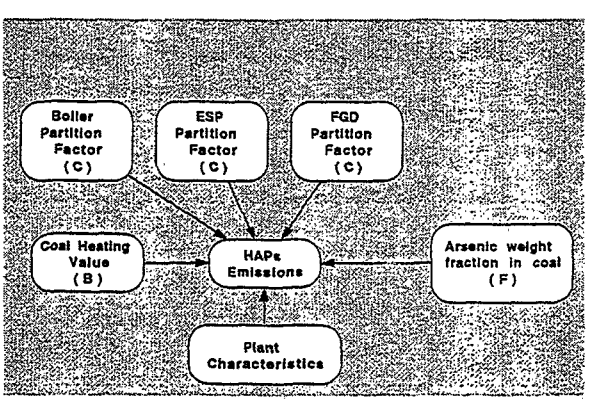

NC STATE

## IMPLEMENTATION IN DEMOS




NC STATE

## Arsenic Emissions Model in Analytica




NC STATE

## Sample Size

- Variability
  - –Based on criteria for accuracy of simulation
  - –Mean
  - –Fractiles
- Uncertainty
  - –Same as variability, or
  - –Tolerance Intervals (e.g., Hoffman and Hammond, 1994)

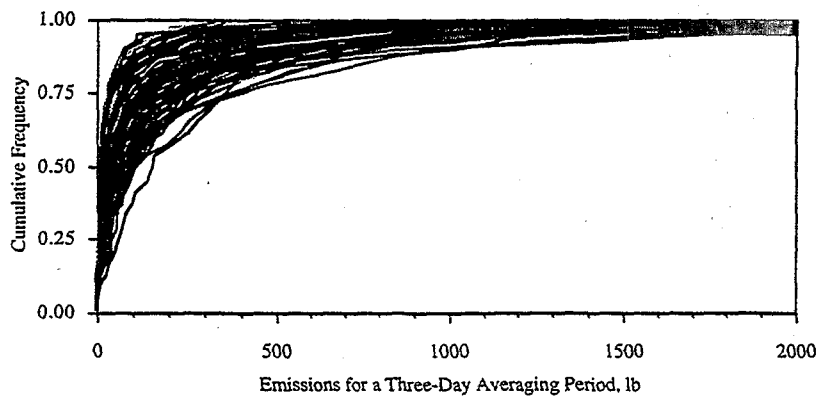
NC STATE

## ANALYSIS AND REPORTING

- Displaying results
- Interpreting results
- Determining key sources of uncertainty and variability

NC STATE

## Results of a 2-D Analysis



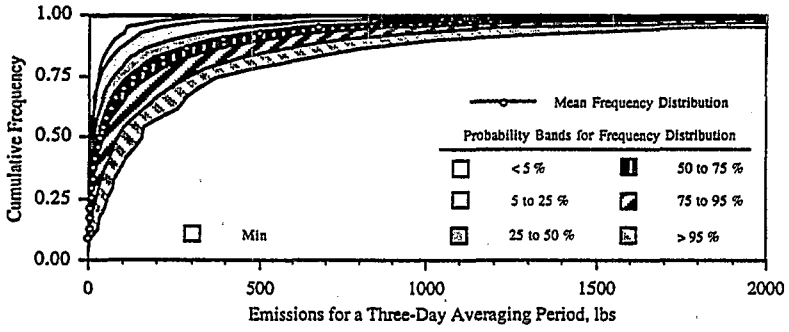

NC STATE

## TYPES OF MODEL RESULTS: Examples

- Mean CDF for variability
- Uncertainty for a specific individual
- Uncertainty for a randomly selected individual
- Uncertainty for any given exposure/ risk level
- Uncertainty for any given fractile of the population

NC STATE

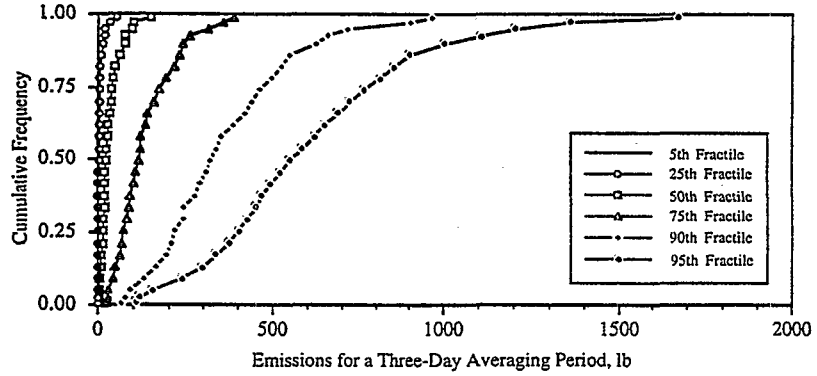
## Probability Bands for Frequency Distribution




*Uncertainty Increases for a Given Fractile*
*Uncertainty for a Given Emission Rate is Bounded (0,1)*

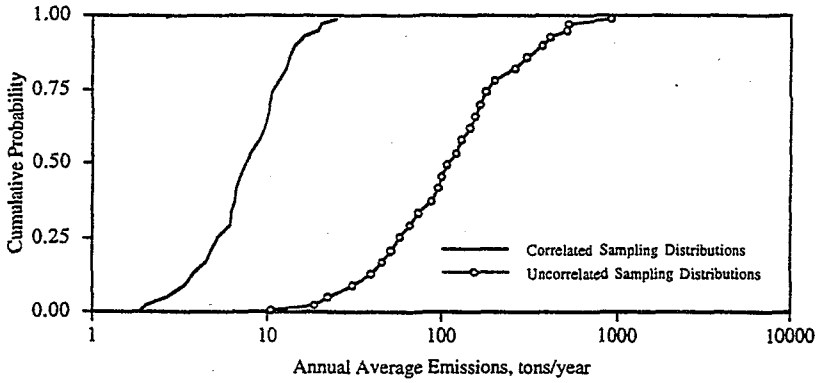
NC STATE

## Uncertainty in Emissions for a Given Fractile




NC STATE

## Importance of Dependence Among Sampling Distributions




*Failure to account for dependence between sampling distributions can produce misleading results*

NC STATE

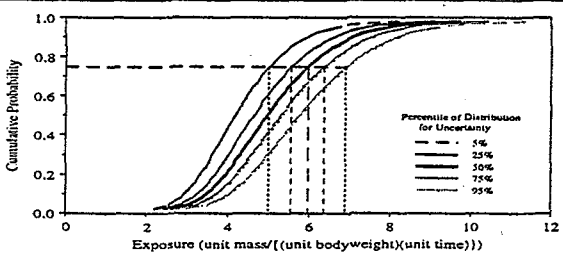
## USING ERROR BARS TO REPRESENT UNCERTAINTY

**Simple Model:**

**E = V + U**

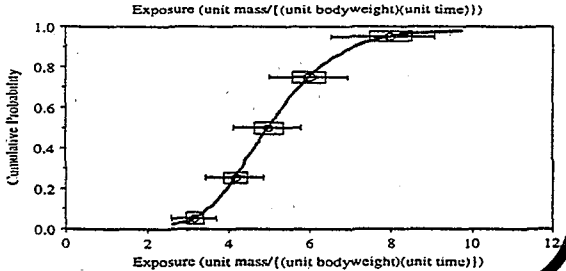
**V ~ LN**

**U ~ N**



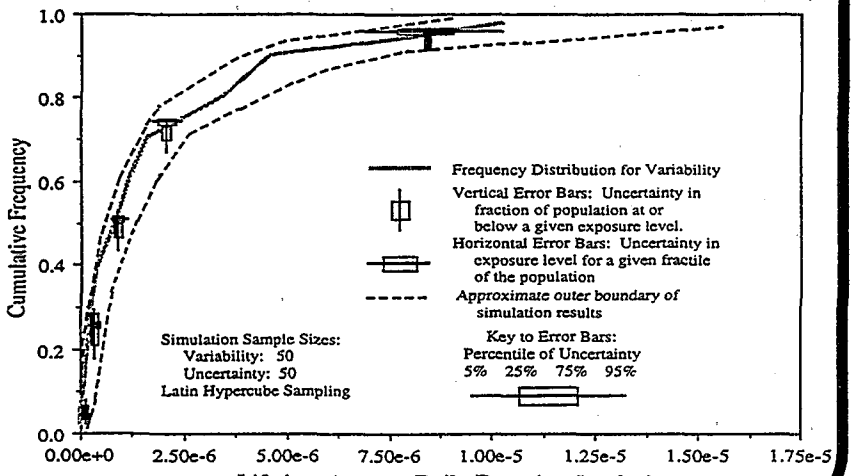

**Error Bar: Percentile of Uncertainty**

5% 25% 75% 95%



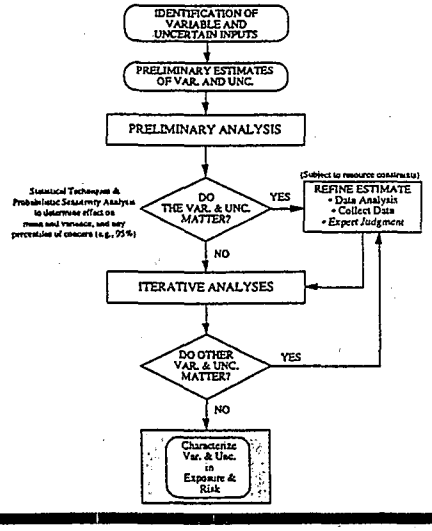

NC STATE

# COMMUNICATING VARIABILITY AND UNCERTAINTY




NC STATE

# SCREENING AND ITERATION




NC STATE

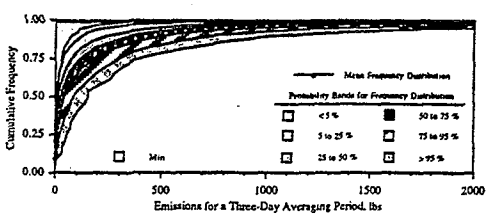
## Identifying Key Model Inputs

- Sensitivity Analysis
  - Change Input Assumptions
  - Compare Model Results
- Statistical Analysis
  - Rank Correlations
  - Partial Rank Correlation Coefficients
  - Standardized Rank Regression Coefficients

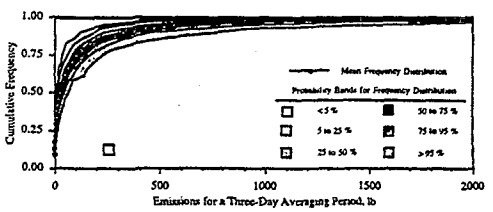
NC STATE

## Probabilistic Sensitivity Analysis

**Base Case**

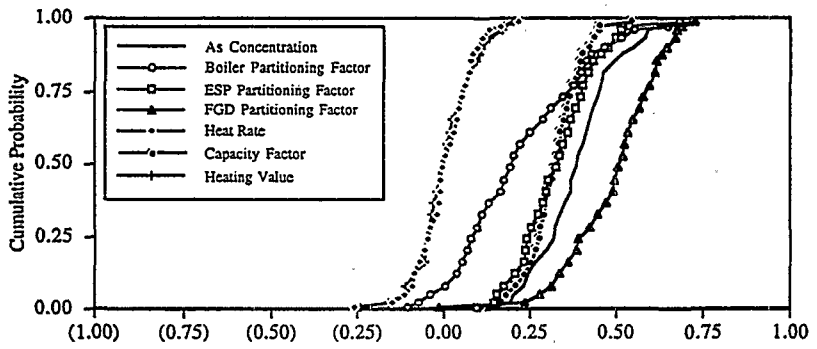



**Without Uncertainty in Partitioning Factors**




NC STATE

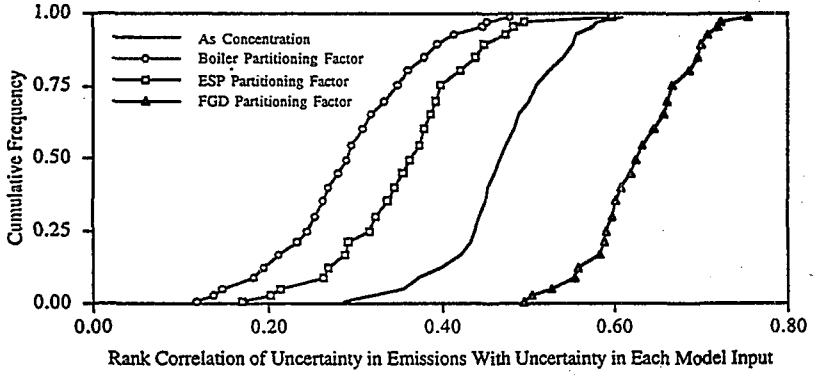
## Uncertainty Regarding Key Sources of Variability




*Ambiguity Regarding Key Sources of Variability is Due to Uncertainty*

NC STATE

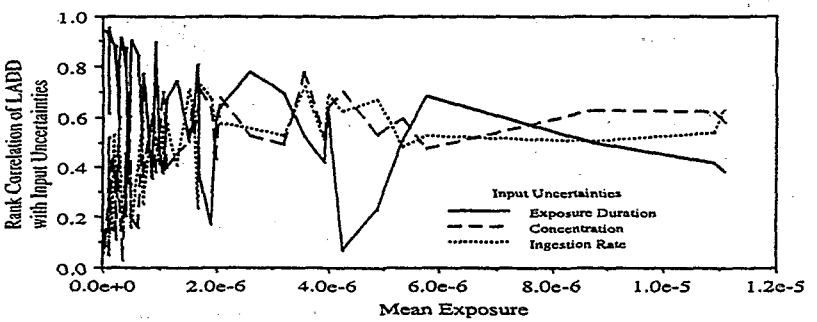
## Variability in Key Sources of Uncertainty




*Key Sources of Uncertainty Differ From One Member of the Population to Another*

NC STATE

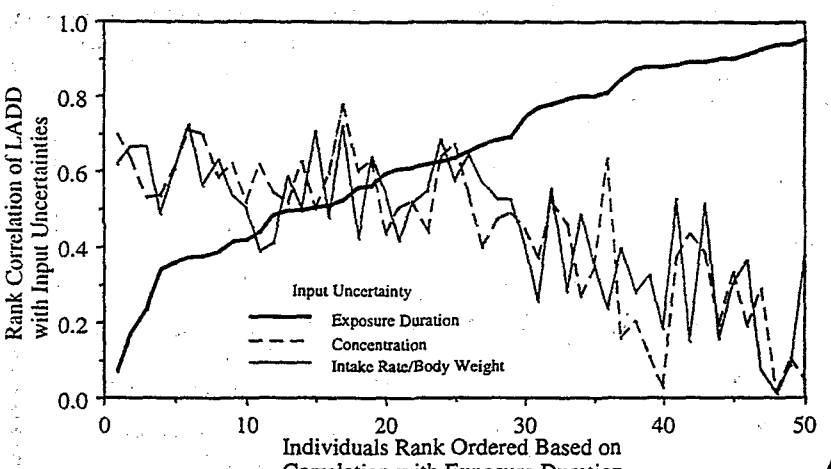
# IDENTIFYING KEY UNCERTAINTIES




- **Can get similar exposures for different reasons**

NC STATE

# IDENTIFYING KEY UNCERTAINTIES




NC STATE

## IMPLICATIONS OF TWO-DIMENSIONAL SIMULATION

- Conceptual difference between variability and uncertainty
- To specify a point estimate, need two coordinates:
  - Percentile of population
  - Percentile of uncertainty

e.g., 95% probability that 90% of population faces an exposure less than X

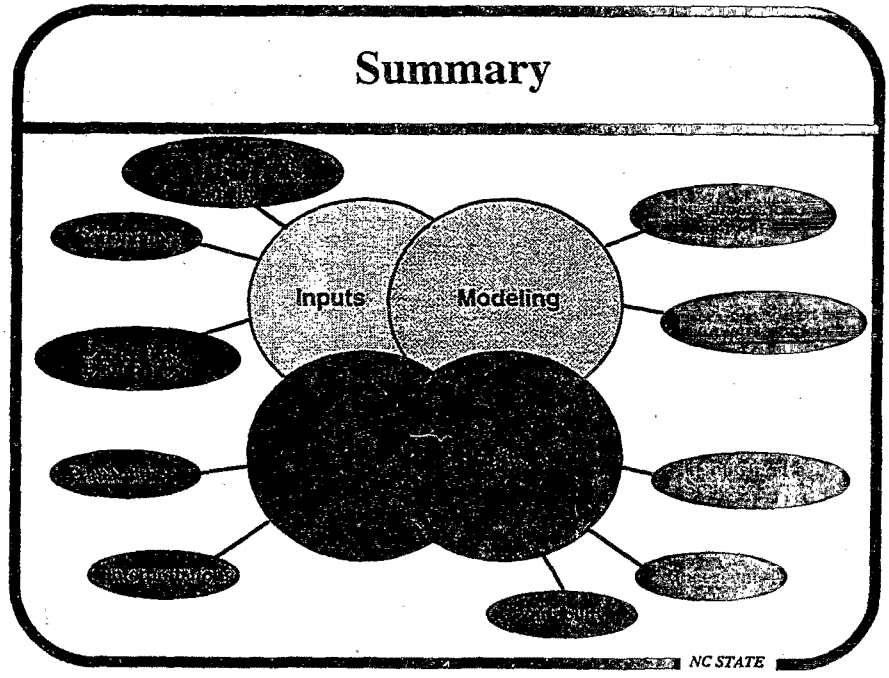
- Can prioritize data collection/research separately for uncertain and variable quantities
- Must be able to disaggregate input data into variable and uncertain components

NC STATE

## DOCUMENTING THE ANALYSIS

- Be clear about scenarios, models and causal assumptions
- Summarize input uncertainty estimates in a table
- Explain each uncertainty estimate
- State sources of information (e.g., data, expert, literature)
- Focus debate: key uncertainties
- Discuss robustness of results to different:
  - Models
  - Correlation structures
  - Expert judgments
  - Uncertainty characterizations
- Discuss, compare qualitative, quantitative uncertainties
- Summarize bottom line and limitations in one place
  - Graphical
  - Laundry list
  - Summary statistics
  - Priorities for future work

NC STATE


Summary
Inputs
Modeling
NC STATE

# Risk Assessment Forum

# Workshop on Monte Carlo Analysis

May 13 - 15, 1996

New York City, New York

**Timothy M. Barry**
**CPAD/OPPE**
**Environmental Protection Agency**

tbarry CPAD/05.08.96

## Distributions on a Budget

*.....dancing as fast as you can*

- ✖ *No Time*
- ✖ *No Money*
- ✖ *No Hope of Gathering Data*
- ✖ *Relentless Scrutiny from Many Quarters (both friendly and not so friendly)*
- ✖ *No Hope of Changing Any of the above*

tbarry CPAD/05.08.96

**Our basic premise was that we could assign each exposure variable to a distributional family whose exact shape may not be well known (uncertain)**

✔ *Each exposure variable may be assigned to a distributional family...*

$$X \sim PDF_V(\alpha,\beta,...) \qquad \textit{e.g.,}\ \ \textit{Concentration} \sim LN(\mu,\sigma)$$

✔ *...whose exact shape, within that distributional family, may be uncertain*

$$\alpha \sim PDF_U(a_1,a_2,...), \qquad \beta \sim PDF_U(b_1,b_2,...) \quad \textit{etc.}$$

tbarry CPAD/05.08.96

# Deciding on the Distributional Family

- ✔ What, if anything, is known about mechanisms giving rise to the variable?
- ✔ What data are available? What is their quality? How representative are the data of the variable in the problem of interest?
- ✔ Is the variable discrete or continuous?
- ✔ What are the bounds of the variable?
- ✔ Is the variable known, or thought to be, skewed or symmetric?

tbarry CPAD/05.08.96

# Overview of Our experience in the Radon in Drinking Water Exposure Assessment

- **Focus on 4 exposure model variables**

- **From "data rich" to "data poor"**

- **Exposure Variables**
  - *radon-222 concentrations in community groundwater delivery systems*
  - *radon-222 water-to air transfer factor*
  - *radon-222 equilibrium factor*
  - *residential occupancy factor*

tbarry CPAD/05.08.96

# Radon Concentrations in Community Groundwater Systems

## Data Sources

- ***National Inorganic and Radionuclides Survey (NIRS)***

...a stratified random sample of the 47,770 community ground water supply systems inventoried in the Federal Reporting Data Systems (FRDS) in 1984.

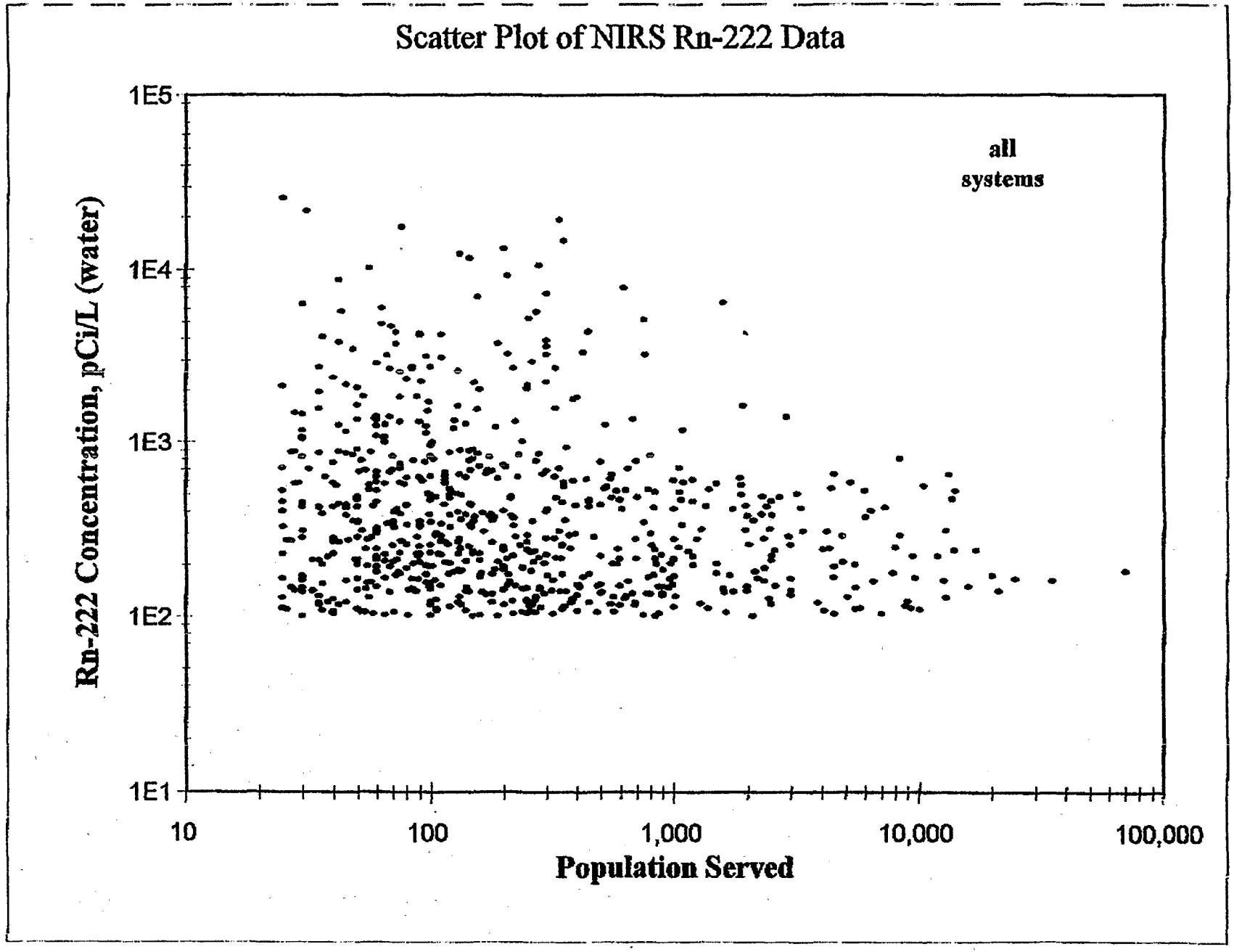
NIRS included 1,000 systems, or approximately 2.1% of the total FRDS inventory. Of the 1,000 systems surveyed, 990 responded. Of the 990 systems which responded, eight were samples were excluded, leaving 982 systems in the data base.

tbarry CPAD/05.08.96

| Ground Water Population Served | Number of FRDS Sites (fiscal 1985) | NIRS Target Sites | Re-Stratified Size Categories (1992) | FRDS Inventory (1992) | NIRS Sites (1992) |
|---|---|---|---|---|---|
| very small 25 - 500 | 34,040 71.4% | 716 71.6% | very, very small 25 - 100 | 16,634 36.5% | 335 34.1% |
| | | | very small 101 - 500 | 15,422 33.8% | 334 34.1% |
| small 501 - 3300 | 10,155 21.3% | 211 21.1% | small 501 - 3300 | 9,952 21.8% | 232 23.6% |
| Medium 3301 - 10,000 | 2,278 4.8% | 47 4.7% | Medium 3301 - 10,000 | 2,302 5.0% | 53 5.4% |
| Large & very Large >10,001 | 1,227 2.6% | 26 2.6% | Large & very Large > 10,001 | 1,316 2.9% | 28 2.9% |
| TOTALS | 47,700 | 1000 100.0% | TOTALS | 45,626 100.0% | 982 100.1% |

Summary characteristics for community groundwater supply systems used for estimating radon occurrence.

tbarry CPAD/05.08.96


Scatter Plot of NIRS Rn-222 Data
all
systems
Rn-222 Concentration, pCi/L (water)
1E5
1E4
1E3
1E2
1E1
10
100
1,000
10,000
100,000
Population Served

# Parameterization of the NIRS Radon Concentration Data

Approximately 28% of the NIRS systems (275 out of 982) had radon concentrations below the minimum reporting level (MRL) - represent Type I left-censored samples

- **Maximum Likelihood Function for Type I, Left-Censored Samples**
  The sample likelihood function for left-censored Type I data with *r* censored samples is

$$L(\theta) = \frac{N!}{r!} \times \left[ \int_0^{x_{r+1}} f(x;\theta)\,dx \right]^r \times \prod_{k=r+1}^{N} f(x_k;\theta) = \frac{N!}{r!} \times \left[ F(x_{r+1};\theta) \right]^r \times \prod_{k=r+1}^{N} f(x_k;\theta)$$

where $N$ is the total sample size, $f(x,\theta)$ is the pdf and $\theta$ is a vector of unknown parameters. The values of $\theta_j$ which are sought are those values which maximize the logarithm of the sample likelihood, subject to constraints on parameters $\theta_j$.

***Generally, numerical methods must be used to obtain a solution set.***

tbarry CPAD/05.08.96

## ❒ Chi-Square (Least Squares) Minimization.

...an alternative for estimating parameters of a model in which the minimum of the quantity

$$\chi^2 = \sum_{k=1}^{N} \left[ \frac{y_k - y(x_k;\theta)}{\sigma_k} \right]^2$$

is sought where $y(x;\theta)$ is the model to be fitted to $N$ data points $(x_k,y_k)$ and $\sigma_k$ is the standard deviation of the *kth* datum. As with the method of maximum likelihood, the solution to equation is found by solving $m$ simultaneous equations

❑ ***PP Regression***. Percentile-percentile regression (i.e., regression on order statistics) is a specific case of $\chi^2$ minimization in which the values of $\theta_j$ which minimize the sum of the squared difference between the predicted and observed cumulative distribution functions are sought, i.e.,

$$MIN: \quad \sum_{k=r+1}^{N} \left[ \frac{k}{N+1} - F(x_k;\theta) \right]^2$$

## Parameter Estimation for Type I Censored Distributions: Special Case of the Lognormal Distribution.

A number of simplifications can be made if the distribution being fit is the lognormal, i.e., $X \sim LN(\mu,\sigma)$.

- **For MLE...**

$$L^* = -n_2 \ln \sigma - \frac{1}{2} \sum_{k=n_1+1}^{N} y_k^2 + n_1 \ln[\Phi(\xi)] + constant$$

where $y_k = [\ln(x_k)-\mu]/\sigma$, $\xi = [\ln(x_{r+1}) -\mu]/\sigma$ and $\Phi(\xi)$ is the normal cumulative distribution.

- **For ROS**...it is possible to express the solution in closed form if regression is against the Gaussian *Z-score*, i.e.,

$$MIN: \quad \sum_{k=n_1+1}^{N} [\ln(x_k) - \hat{u} - \hat{\sigma} Z_k)]^2 \qquad where \quad Z_k = \Phi^{-1}\left(\frac{k}{N+1}\right)$$

## Distributions Tested

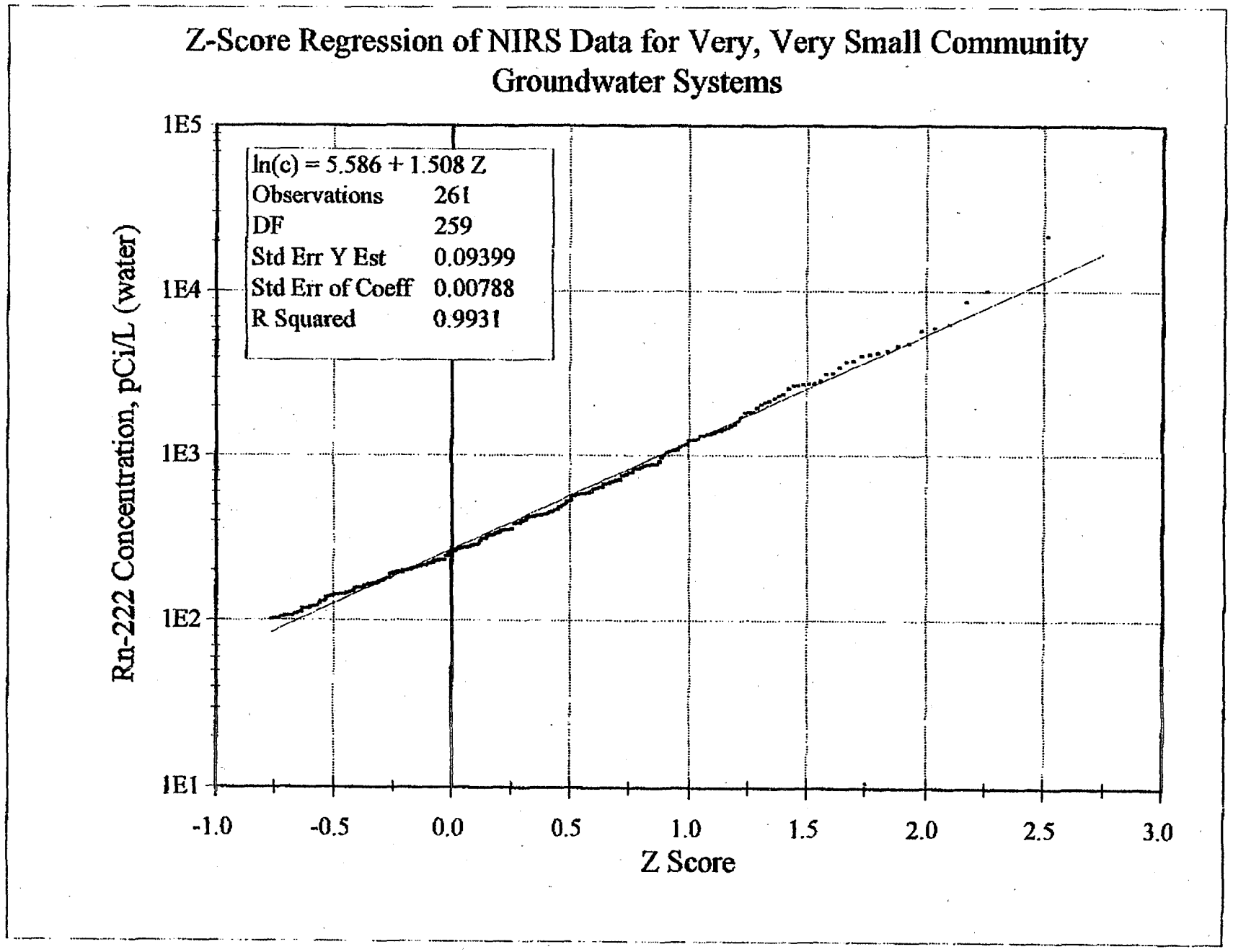
- Graphical exploration of the data showed it to be right-skewed, with a tails spanning up to 2½ orders of magnitude

- The radon concentration data were fitted to five skewed distributions, using both MLE and ROS parameter estimation techniques
  - *lognormal*
  - *gamma & log-gamma*
  - *Weibull & log-Weibull*
  - *Pearson Type 5 & log-Pearson 5*
  - *Pearson Type 6 & log-Pearson 6*

- Generally, the two-parameter gamma fit best, but the data were also well fit by the lognormal, except for small systematic deviations in the upper tail.

- Because of the advantages associated with lognormal pdfs and the marginal gains of the gamma, radon concentrations were models as lognormal.

tbarry CPAD/05.08.96


Z-Score Regression of NIRS Data for Very, Very Small Community Groundwater Systems
ln(c) = 5.586 + 1.508 Z
Observations 261
DF 259
Std Err Y Est 0.09399
Std Err of Coeff 0.00788
R Squared 0.9931
Rn-222 Concentration, pCi/L (water)
1E5
1E4
1E3
1E2
1E1
-1.0
-0.5
0.0
0.5
1.0
1.5
2.0
2.5
3.0
Z Score

# Characterization of Parameter Uncertainty - the Quality Factor (QF) Approach

- Under the model $C \sim LN(\mu,\sigma)$, how well are $(\mu,\sigma)$ known?

- We ***assumed*** that the NIRS data ***were representative***, and that the only significant uncertainties were attributable to sample size within each stratum.

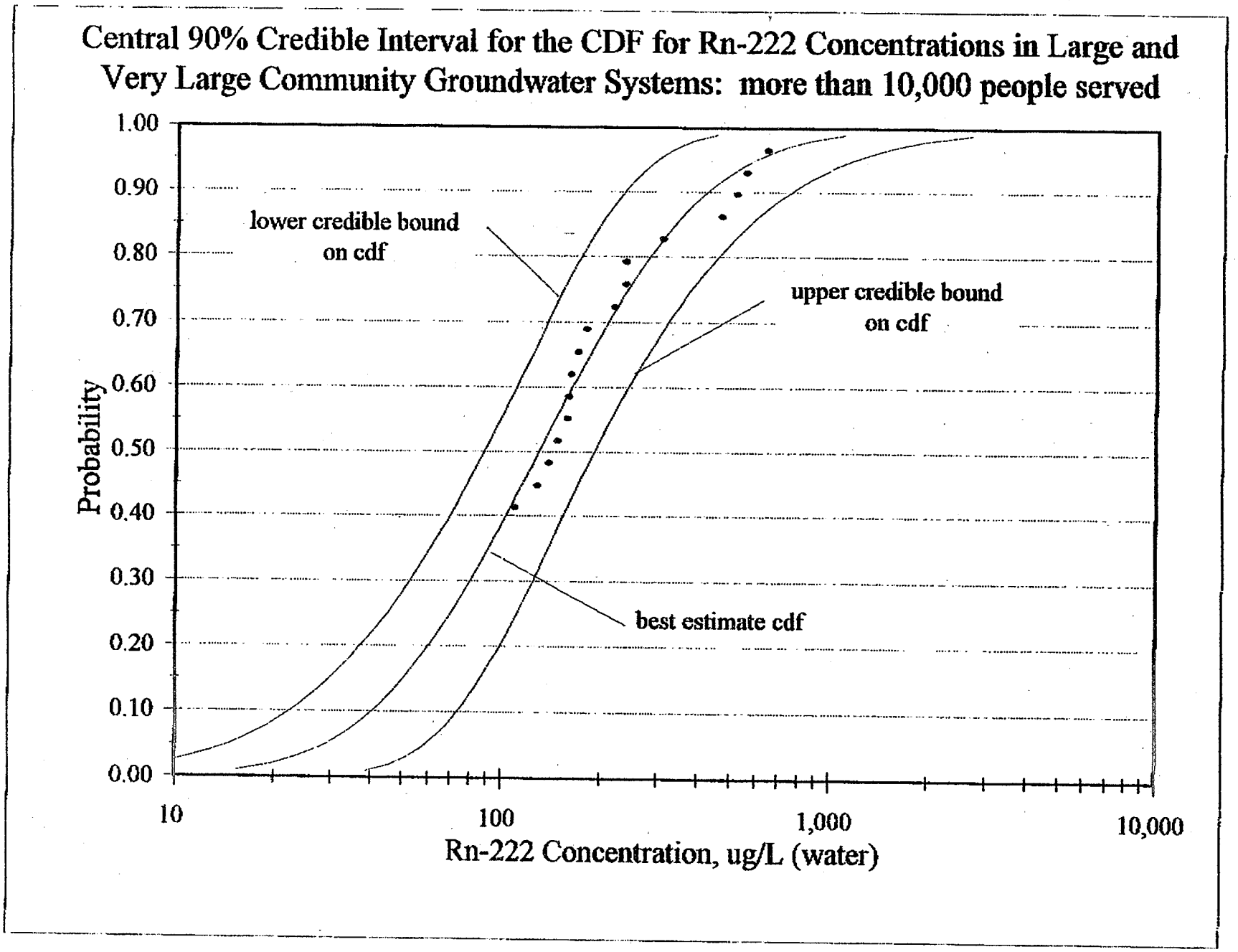
- Sample size uncertainty in the parameters of the radon lognormal distributions were modeled, from classical statistics, as *t* and inverse chi-square distributions, i.e.,

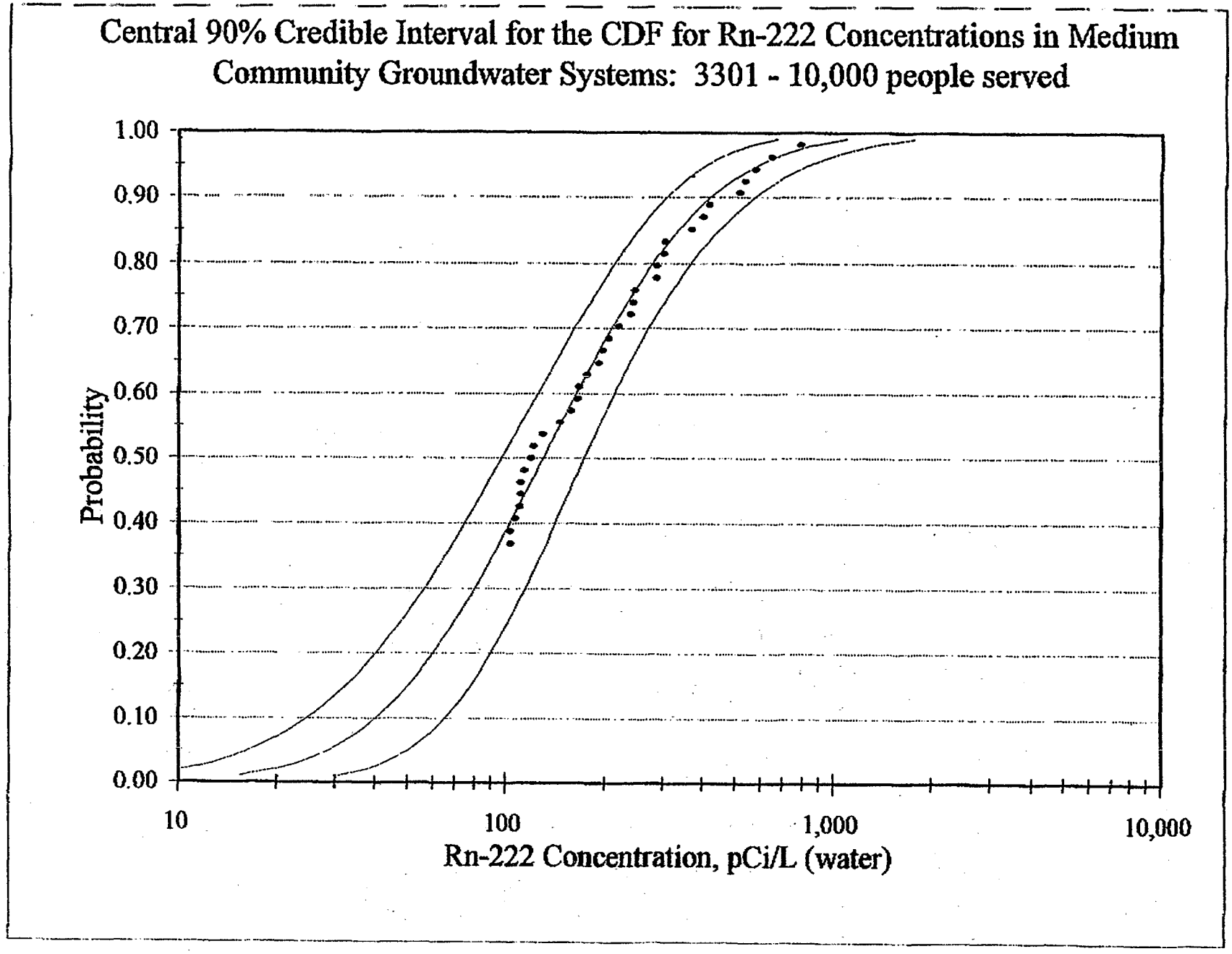
$$\frac{m-\mu}{s/\sqrt{n}} \sim t_{n-1} \qquad \textit{and} \qquad \frac{(n-1)s^2}{\sigma^2} \sim \chi^2_{n-1}$$

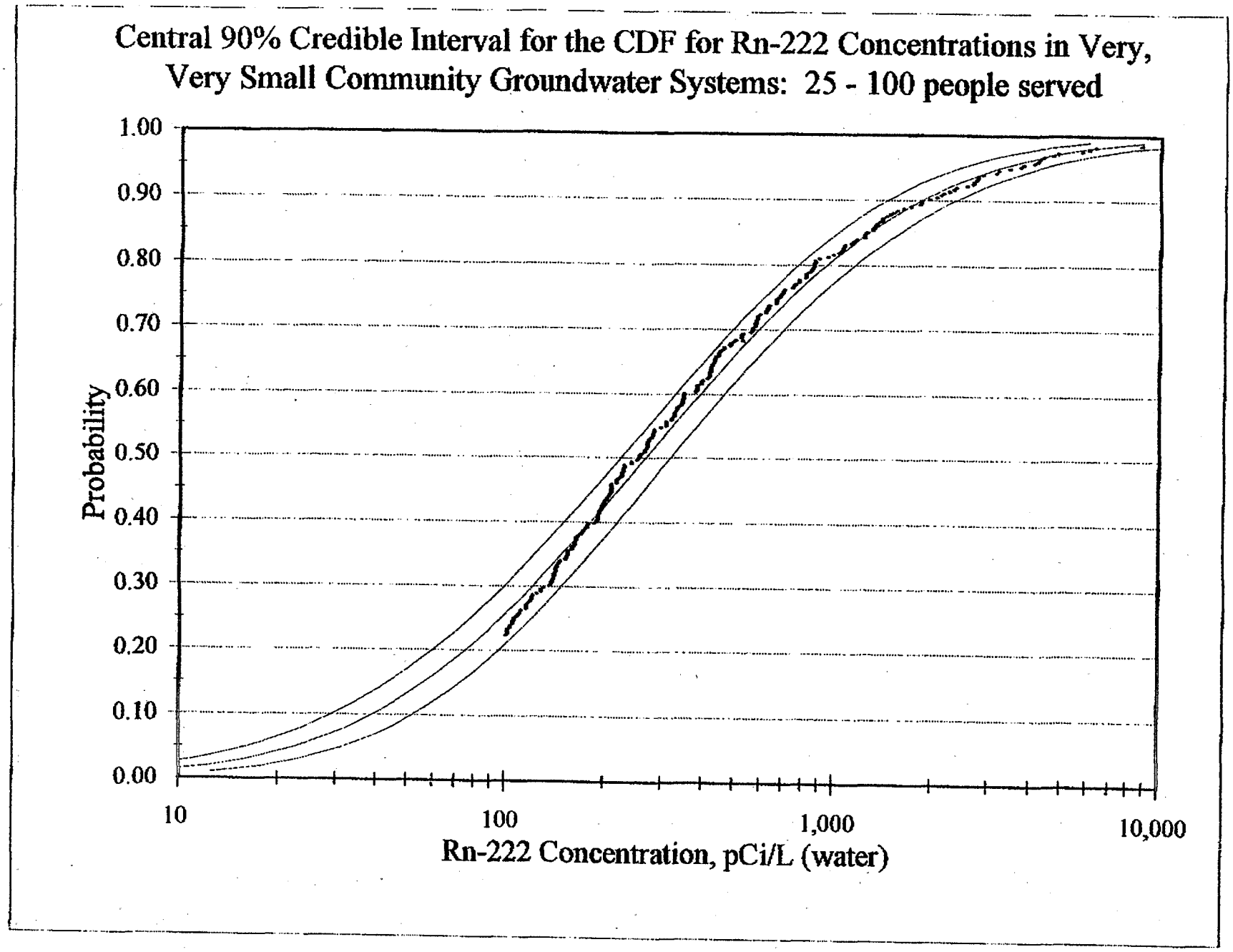
- In specifying our uncertainty in $(\mu,\sigma)$, we invoked a subjective **Quality Factor (QF)** which represented an adjustment to the sample size to reflect the number of samples and well how well we felt the data represented the population of interest.

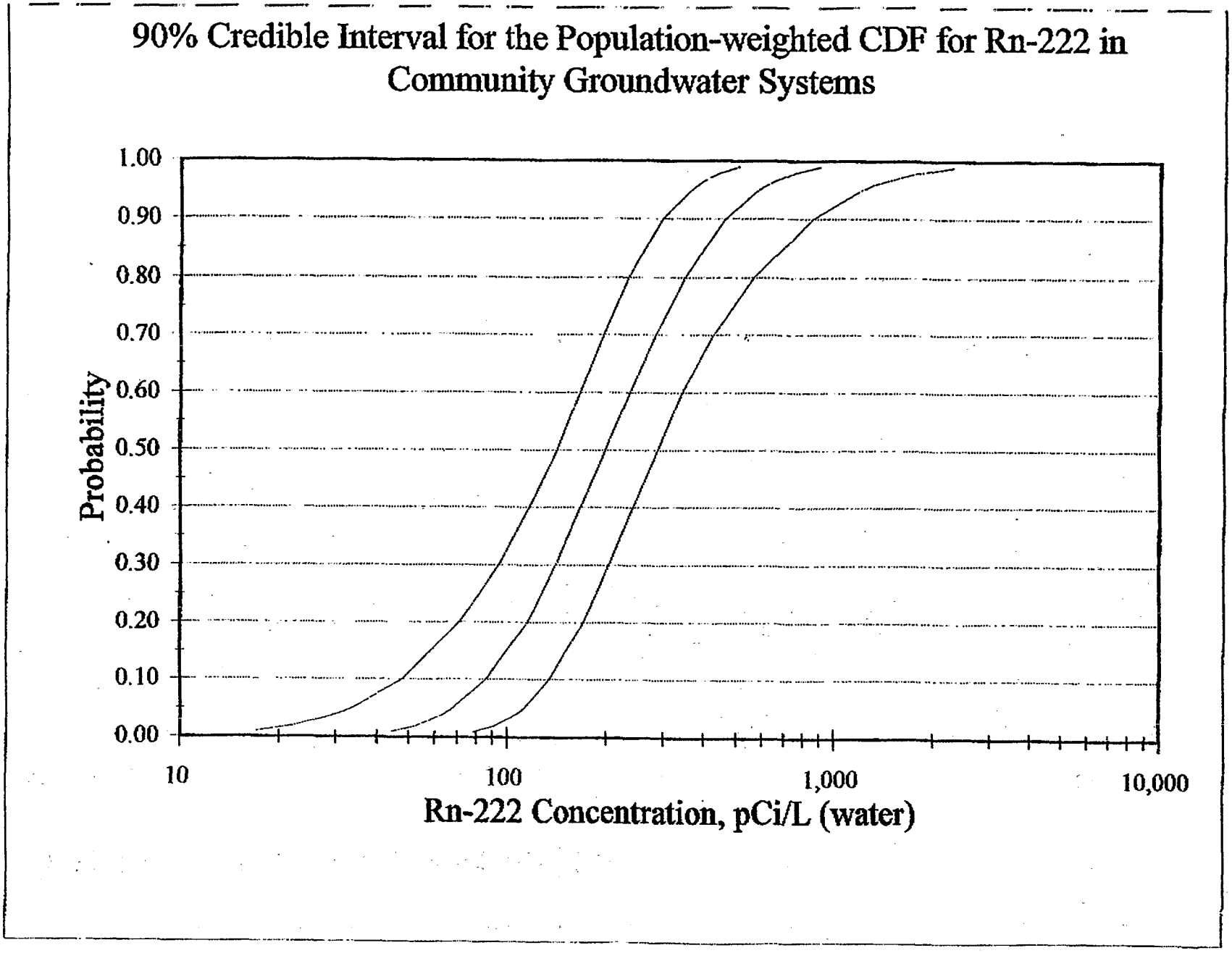
  - $n \rightarrow 10$ = QF small studies and/or not fully representative of population
  - $n \rightarrow 25$ = QF intermediate size, somewhat less than fully representative
  - $n \rightarrow 100$ = QF large and/or acceptably representative

tbarry CPAD/05.08.96


Central 90% Credible Interval for the CDF for Rn-222 Concentrations in Large and Very Large Community Groundwater Systems: more than 10,000 people served
1.00
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
Probability
lower credible bound on cdf
upper credible bound on cdf
best estimate cdf
10
100
1,000
10,000
Rn-222 Concentration, ug/L (water)


Central 90% Credible Interval for the CDF for Rn-222 Concentrations in Medium Community Groundwater Systems: 3301 - 10,000 people served
Probability
1.00
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
10
100
1,000
10,000
Rn-222 Concentration, pCi/L (water)


Central 90% Credible Interval for the CDF for Rn-222 Concentrations in Very, Very Small Community Groundwater Systems: 25 - 100 people served
1.00
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
Probability
10
100
1,000
10,000
Rn-222 Concentration, pCi/L (water)


90% Credible Interval for the Population-weighted CDF for Rn-222 in Community Groundwater Systems
Probability
1.00
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
10
100
1,000
10,000
Rn-222 Concentration, pCi/L (water)

# Water-to-Air Transfer Factor (TF)

- For a simple box model, the steady-state average radon concentration in air attributable to releases from household water is described in terms of the transfer factor

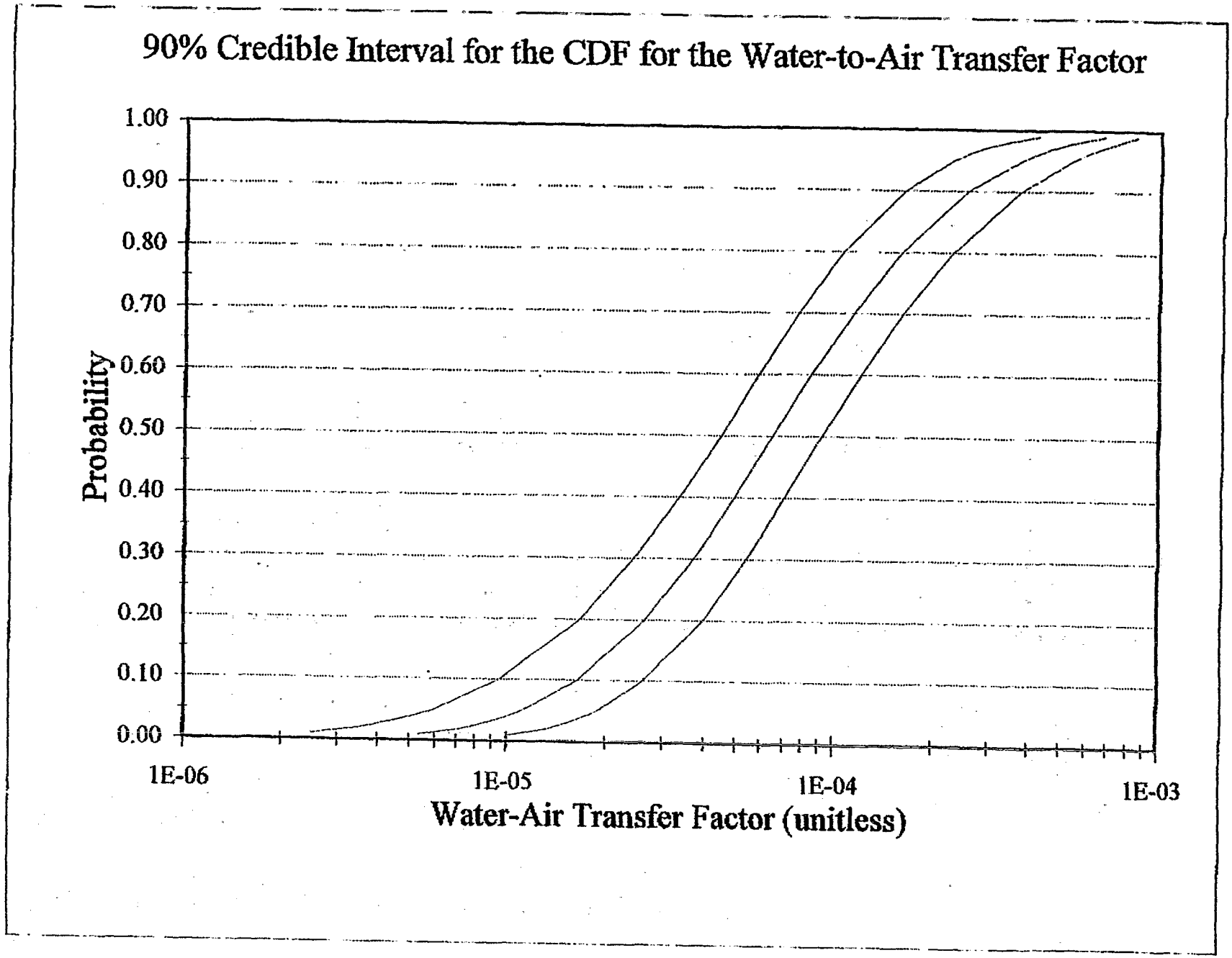
$$TF\,(unitless) = \frac{[Rn\text{-}222]\ (pCi/L\ air)}{[Rn\text{-}222]\ (pCi/L\ water)} = \frac{W \cdot e}{V \cdot \lambda}$$

where W is the whole-house water use rate, use-weighted fractional release rate from water to air, V is the hose volume, and $\lambda$ is the whole-house ventilation rate.

- Nazaroff (1987) used published data on each of these four variables and found that the transfer factor was well described by a lognormal distribution with a geometric mean of 6.57E-05 and a geometric mean of 2.88.

- We used Nazaroff's lognormal distribution with QF = 25 to capture uncertainty in the parameters of the lognormal pdf.

tbarry CPAD/05.08.96


90% Credible Interval for the CDF for the Water-to-Air Transfer Factor
1.00
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
Probability
1E-06
1E-05
1E-04
1E-03
Water-Air Transfer Factor (unitless)

# Uncertain Shapes - Equilibrium Factor

- **The Uncertain Beta Distribution B ~ ($\alpha_1$,$\alpha_2$;min,max)**

$$\alpha_1 = \frac{(mean - min)\cdot(2\cdot mode - min - max)}{(mode - mean)\cdot(max - min)} \qquad \alpha_2 = \alpha_1 \cdot \left( \frac{max - mean}{mean - min} \right)$$

In some cases for bounded variables, we found estimates of ranges and means, but not modes. We did not feel justified in making assumptions about the mode except that it lie between the minimum and the mean if the pdf is right-skewed or between the mean and maximum if the pdf is left-skewed.

For any particular estimate of the mean we treated the pdf as having an unknown shape that could range from nearly flat to highly skewed.

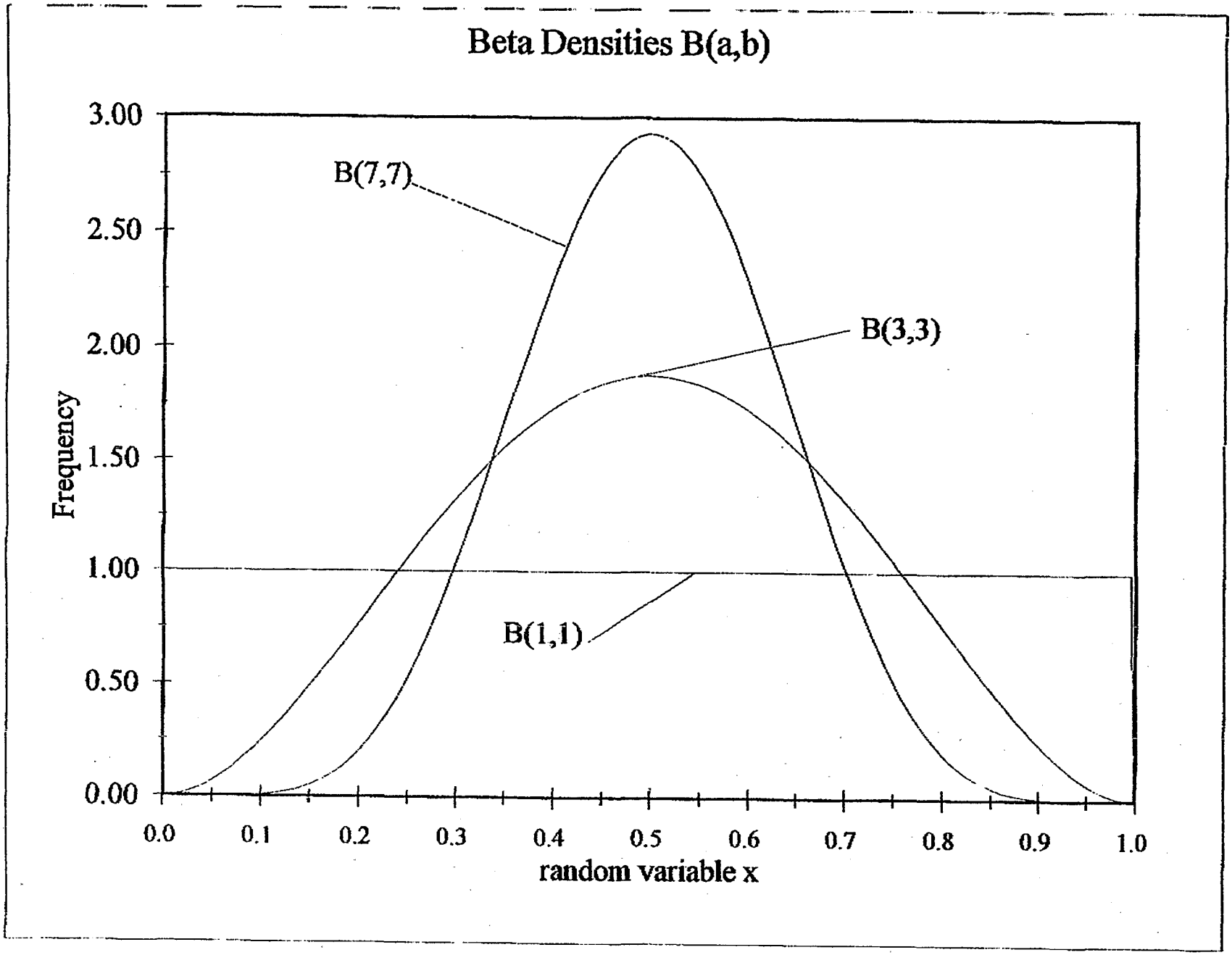
$$PDF_V(EF) \sim B(mode,mean,min,max) \qquad 0.10 \le EF \le 0.90$$

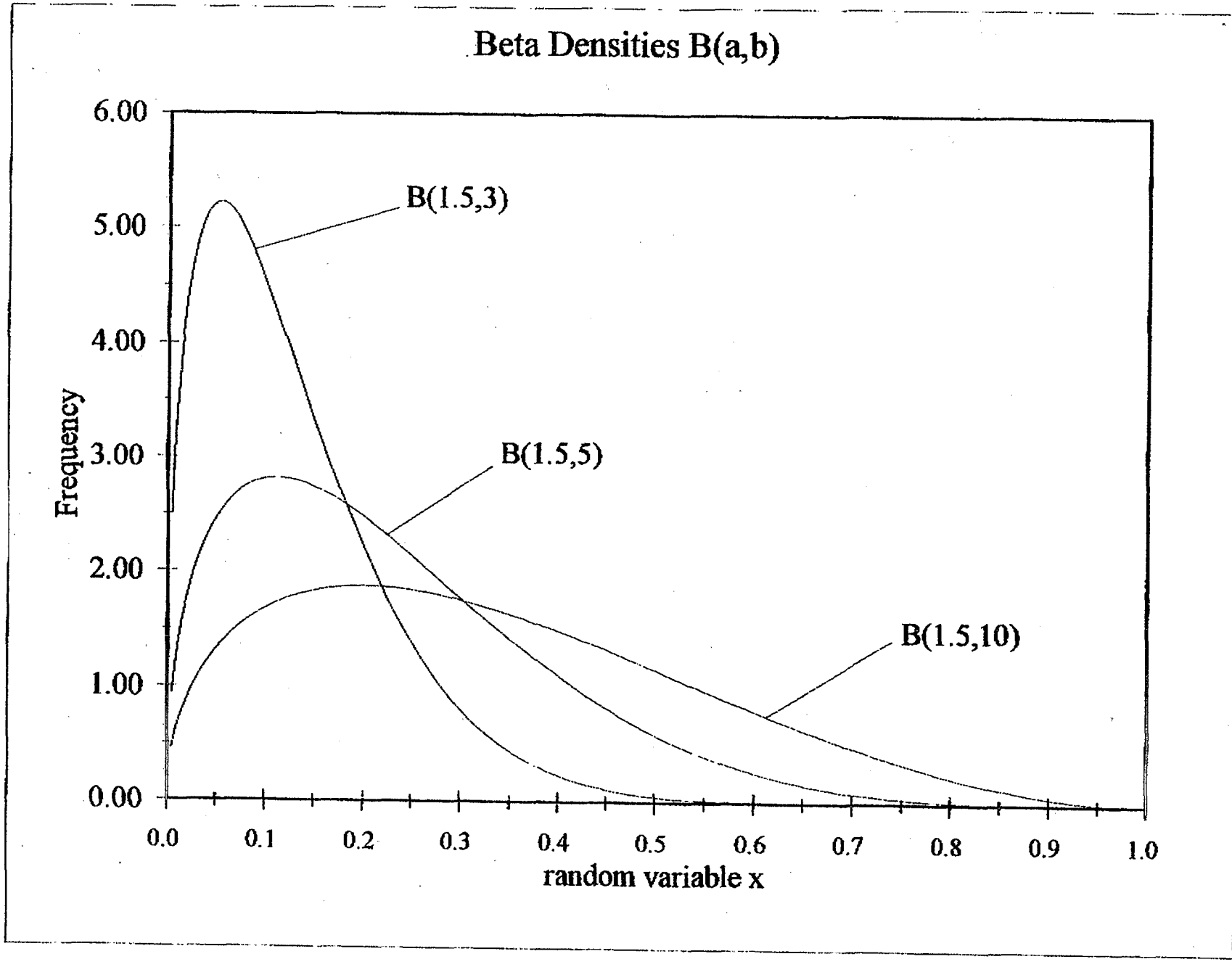
$$PDF_U(mean) \sim U(0.35,0.55)$$

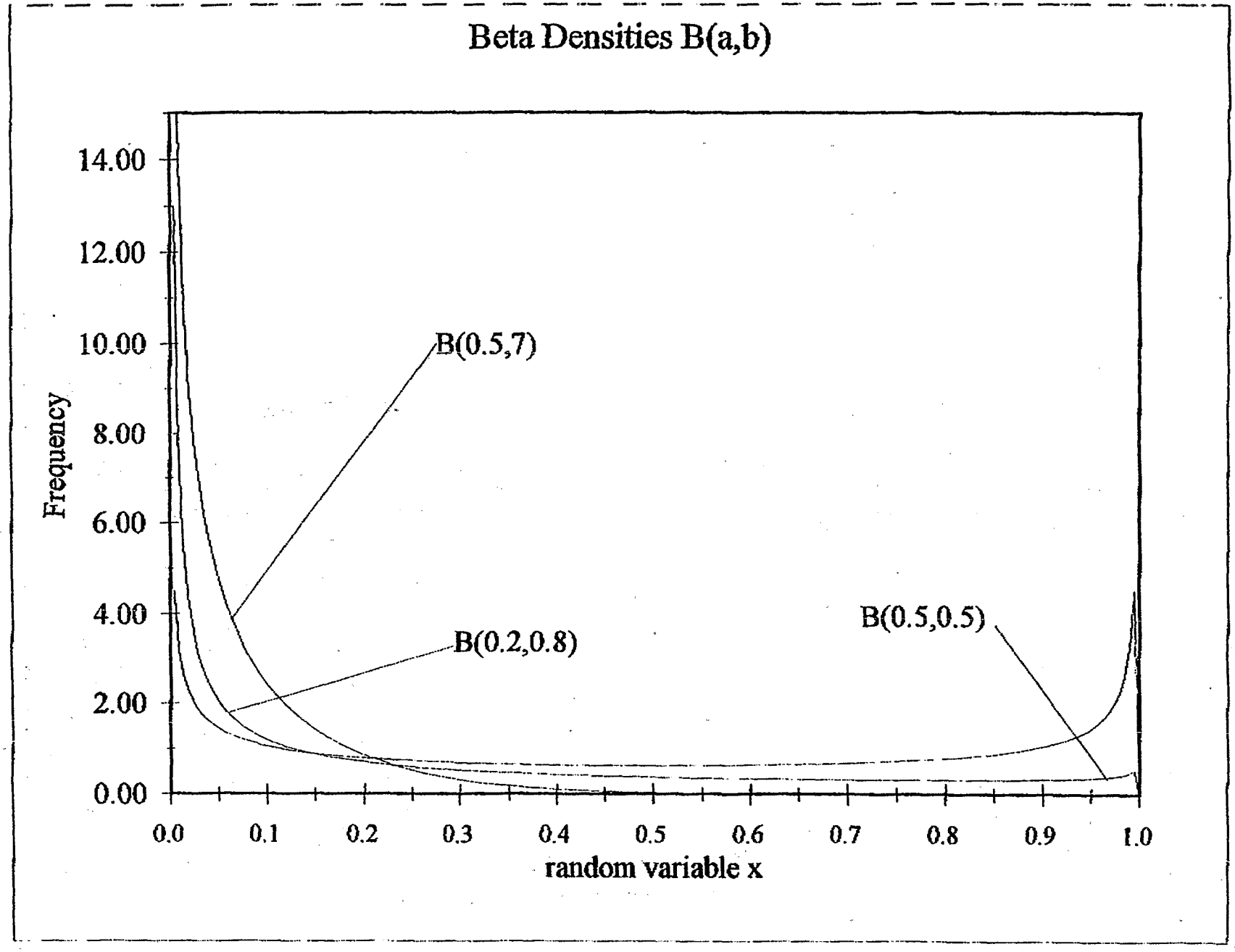
$$PDF_U(mode) \sim U(min,mean) \qquad mean \le \frac{1}{2}(min + max)$$

$$PDF_U(mode) \sim U(mean,max) \qquad mean \le \frac{1}{2}(min + max)$$

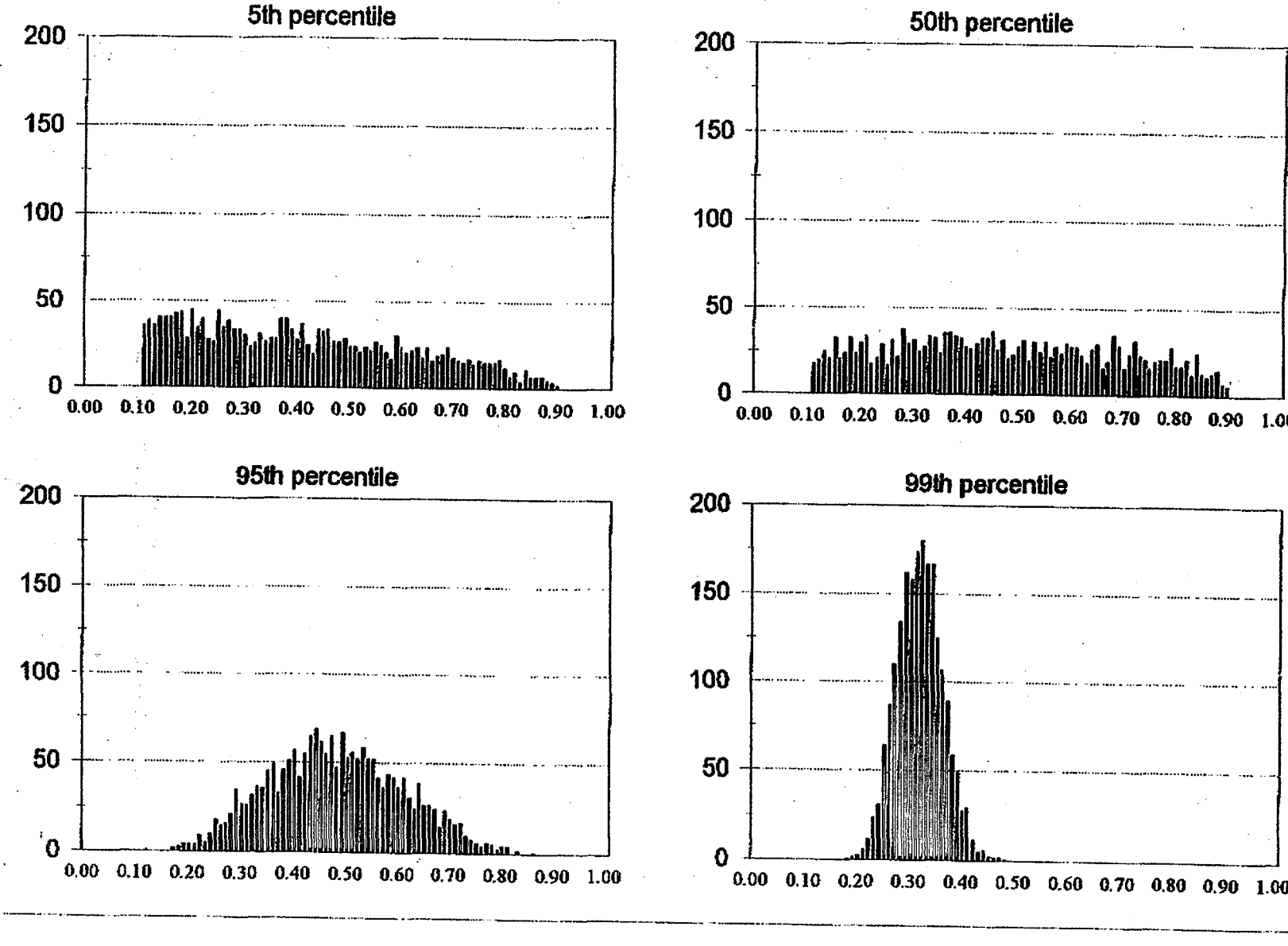
tbarry CPAD/05.08.96

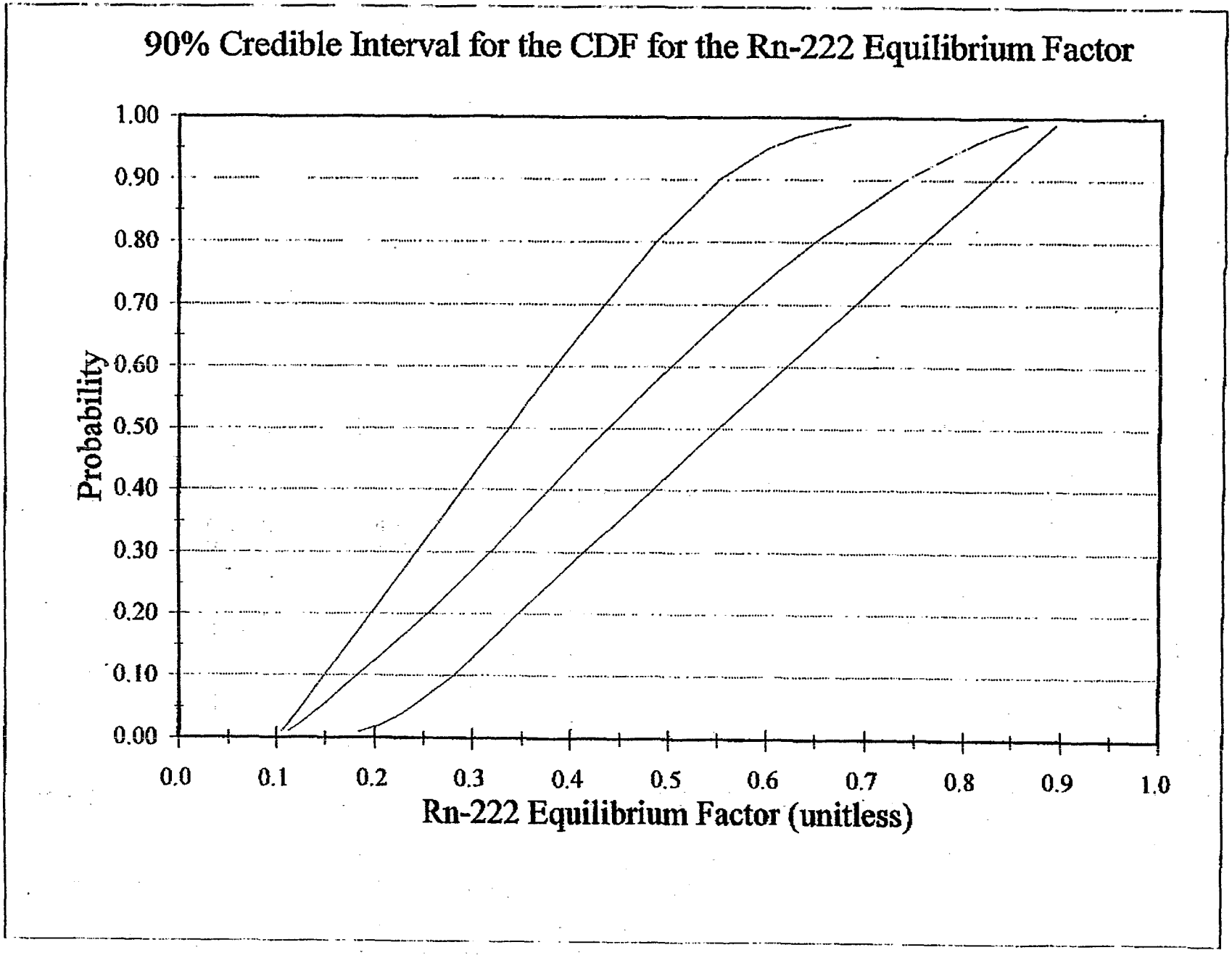

Beta Densities B(a,b)
3.00
2.50
2.00
1.50
1.00
0.50
0.00
Frequency
B(7,7)
B(3,3)
B(1,1)
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
random variable x


Beta Densities B(a,b)
6.00
5.00
4.00
3.00
2.00
1.00
0.00
Frequency
B(1.5,3)
B(1.5,5)
B(1.5,10)
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
random variable x

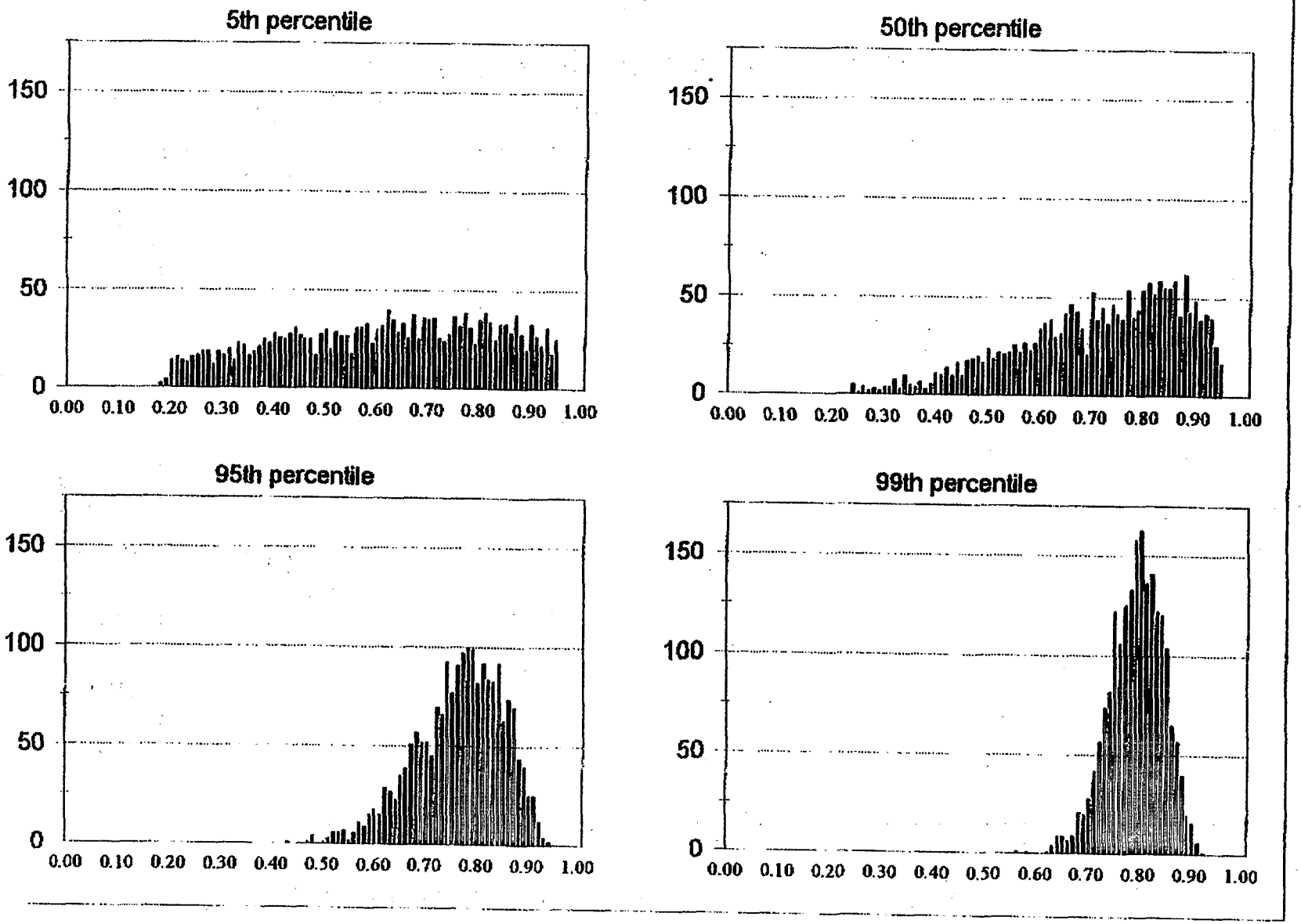

Beta Densities B(a,b)
14.00
12.00
10.00
8.00
6.00
4.00
2.00
0.00
Frequency
B(0.5,7)
B(0.2,0.8)
B(0.5,0.5)
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
random variable x

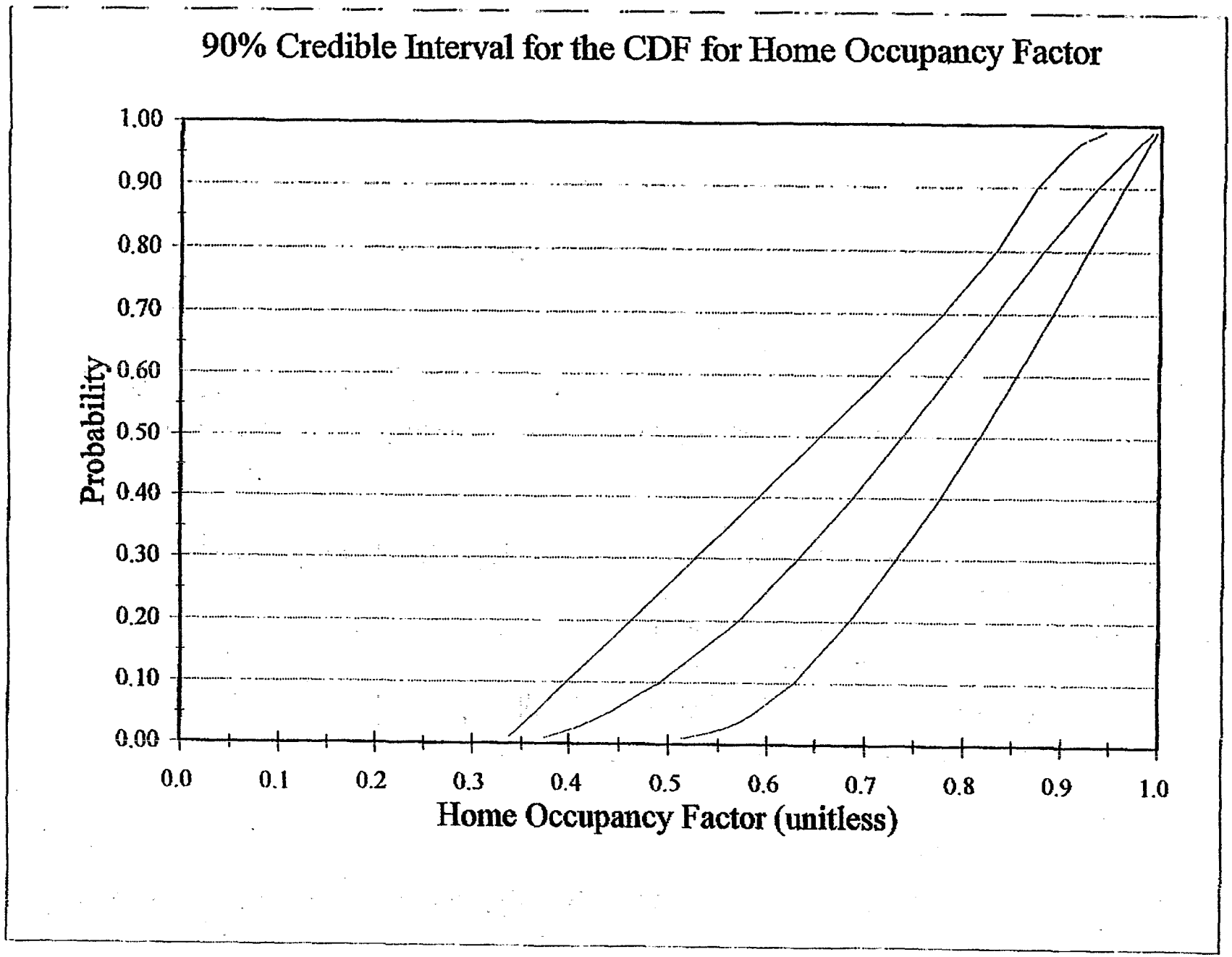
# Range of Allowable Shapes for Equilibrium Factor




90% Credible Interval for the CDF for the Rn-222 Equilibrium Factor
Probability
1.00
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
Rn-222 Equilibrium Factor (unitless)

# Range of Allowed Shapes for Occupancy Factor




90% Credible Interval for the CDF for Home Occupancy Factor
Probability
1.00
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
Home Occupancy Factor (unitless)

# 2-Dimensional Simulation Methodology

$$Exposure = g(X_1, X_2, \ldots) = \frac{\left(\begin{array}{c}Environ\\Concen\end{array}\right) \cdot \left(\begin{array}{c}Contact\\Rate\end{array}\right) \cdot \left(\begin{array}{c}Exposure\\Frequency\end{array}\right) \cdot \left(\begin{array}{c}Exposure\\Duration\end{array}\right)}{\left(\begin{array}{c}Body\\Weight\end{array}\right) \cdot \left(\begin{array}{c}Averaging\\Time\end{array}\right)}$$

$$where\ X_k \sim PDF_V(x_k; \theta_1, \theta_2, \ldots) \qquad and \qquad \theta_k \sim PDF_U(\theta_k; a_k, b_k, \ldots)$$

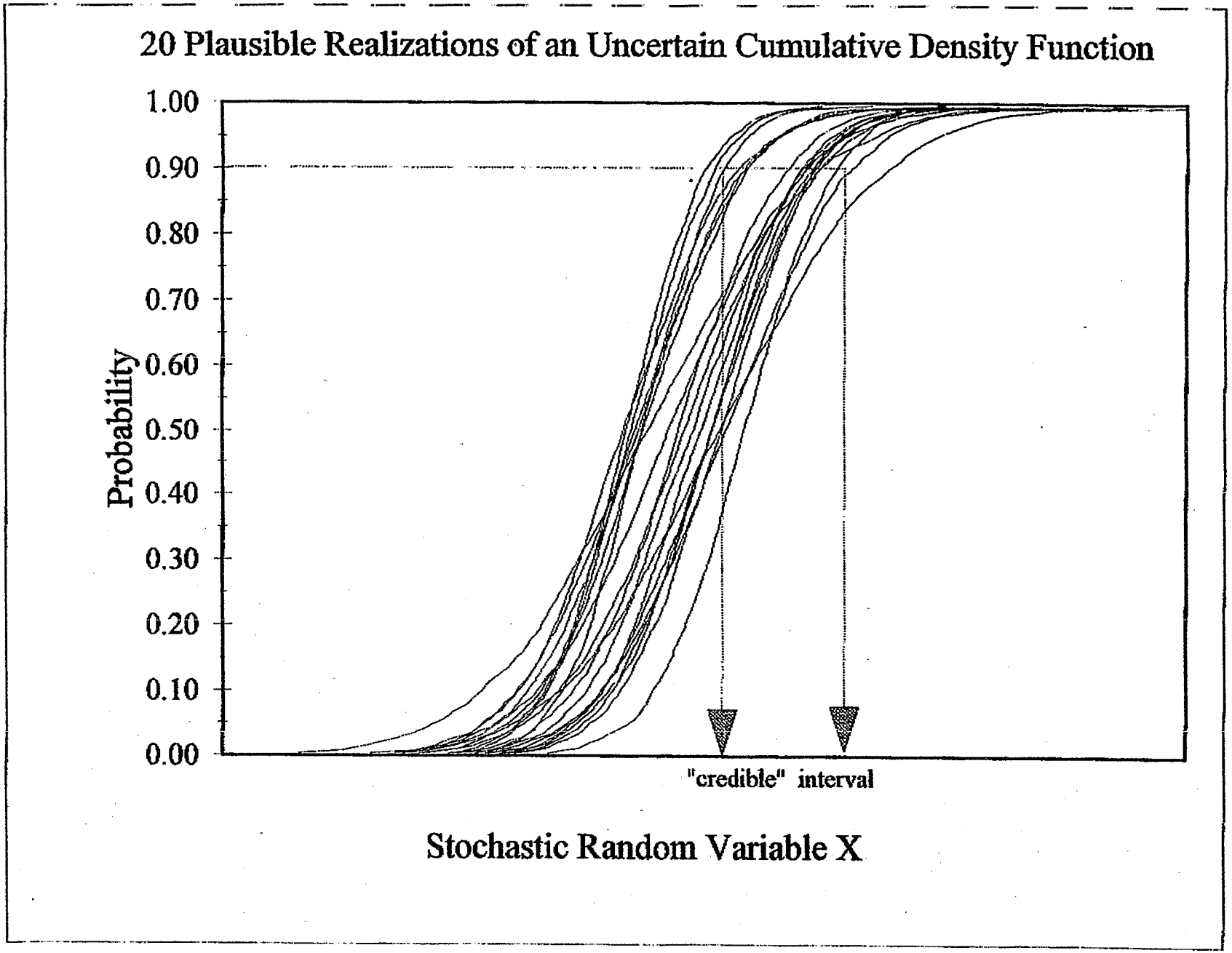
1. For each stochastic variable, randomly pick a plausible parameter set according to $\theta_k \sim PDF_U(\theta_k; a_k\ b_k, \ldots)$

2. Given the parameter set, perform a full Monte Carlo simulation.

3 Record all appropriate and relevant statistics of the input and output variables

4. Repeat steps 1-3 many times, accumulating the range and distribution of relevant input and output statistics

tbarry CPAD/05.08.96


20 Plausible Realizations of an Uncertain Cumulative Density Function
1.00
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
Probability
"credible" interval
Stochastic Random Variable X

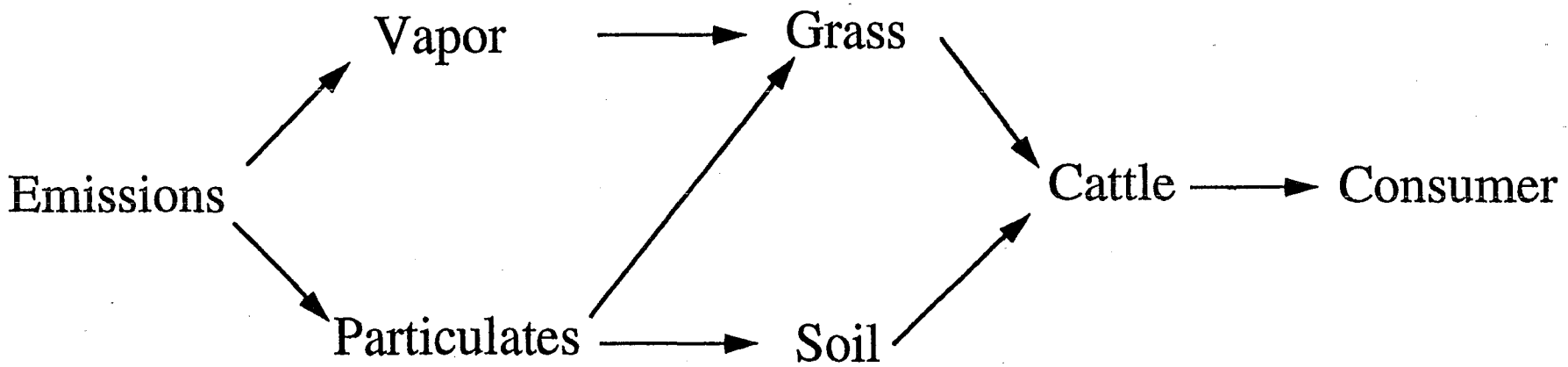
# CASE STUDY: UNCERTAINTY AND VARIABILITY IN INDIRECT EXPOSURES TO TCDD EMITTED FROM A HAZARDOUS WASTE INCINERATOR

**Paul S. Price**
**ChemRisk**

# Case Study: Assessment of Indirect Exposure to TCDD Released from Incinerators Conducted within the Framework of EPA Guidance

- Indirect exposure occurs through the consumption of animal or vegetable foodstuffs that have been contaminated from TCDD deposition and vapor partitioning
- A major indirect exposure pathway is the consumption of beef from cows that graze on impacted pasture
- Both direct deposition and vapor adsorption on plants are important
- Limited time and resources

# Key Pathways for TCDD



# Uncertainty and Variability

- The imprecision of exposure estimates are due to uncertainty and interindividual variability
- Uncertainty - One "true" number that is uncertain because of insufficient knowledge e.g., biotransfer factors or predicted vapor concentration at a specific location
- Variability - A distribution of "true" values that characterizes the variability in nature e.g., amount of beef ranchers consume or location of the ranch

# Method

- Fate and transport of TCDD and human exposure based primarily on EPA (1990,1993) indirect exposure assessment guidance documents

- Hypothetical incinerator located in the Gulf Coast Region of Texas

- Modification of models for TCDD-specific and site-specific information for the hypothetical incinerator

# Approach

- Identify data reasonably available to operator
- Divide variables into three groups
  - point estimates
  - variation dominated
  - uncertainty dominated
- Use a nested loop approach
  - outer loop values for uncertainty
  - inner loop interindividual variation

# Key Components

- Air Emissions
- Air Transport
- Soil
- Beef
- Location
- Consumer

# Air Emissions

- The average emission rate of TCDD under normal operating conditions is well characterized
- Fraction of emissions as vapor is modeled
- Historical data is available on the fraction of time the plant is operational

# Air Transport

- COMPDEP models of vapor and particulates were used
- No loss of vapor to surfaces or particulates
- No consideration of photodegradation of TCDD
- Uncertainty in air modeling results based on historical studies

# Soil and Pasture

- Soil
  - direct deposition
  - long-term equilibrium
- Pasture
  - deposition
  - wash off
  - partitioning based on TCDD-specific data in alfalfa uncertainty is based on laboratory uncertainty

# Beef Concentration

- Based on uptake of pasture and soil
- Biotransfer factors based on Jensen et al. (1981)
- Pasture raised (no supplemental feed)
- Depuration during feed lot time

# Location

- Divide air model receptor grid to equal size blocks representing cattle ranches
- One-half set aside for non-ranching activities
- Air concentrations of TCDD vapor and TCDD deposition rates (wet and dry) are calculated for each ranching block

# Consumers

- Beef sold to the general population not considered
- Consumption of "home grown" beef by ranchers
- Beef consumption rates (variation)
- Duration (variation)
- Fraction home raised (variation)

# Beef Consumption Rates

- Data on variation in short-term consumption rates available on a national basis
- Significant temporal trends in the consumption of beef and beef fat
- Subpopulation of ranchers may have beef intake rates that differ from the national values
- Only short-term variation data available
- Fraction "home raised" based on older study only a point estimate available
- Excellent data theoretically available from USDA

# Source of Data

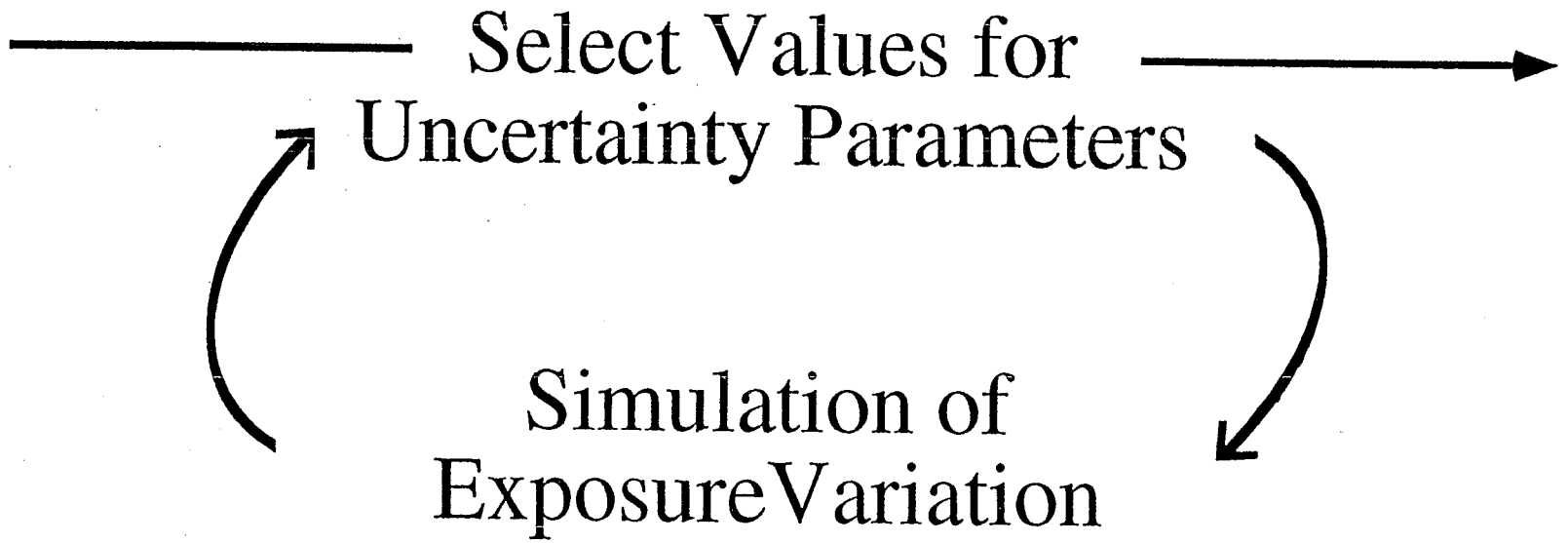
- Site-specific meteorology and topography data
- Site-specific cattle practices
- Facility information (e.g., stack height, etc.) taken from typical values for large hazardous waste incinerators
- Distributions taken from EPA publications wherever possible

# Variables Dominated by Uncertainty or Variability

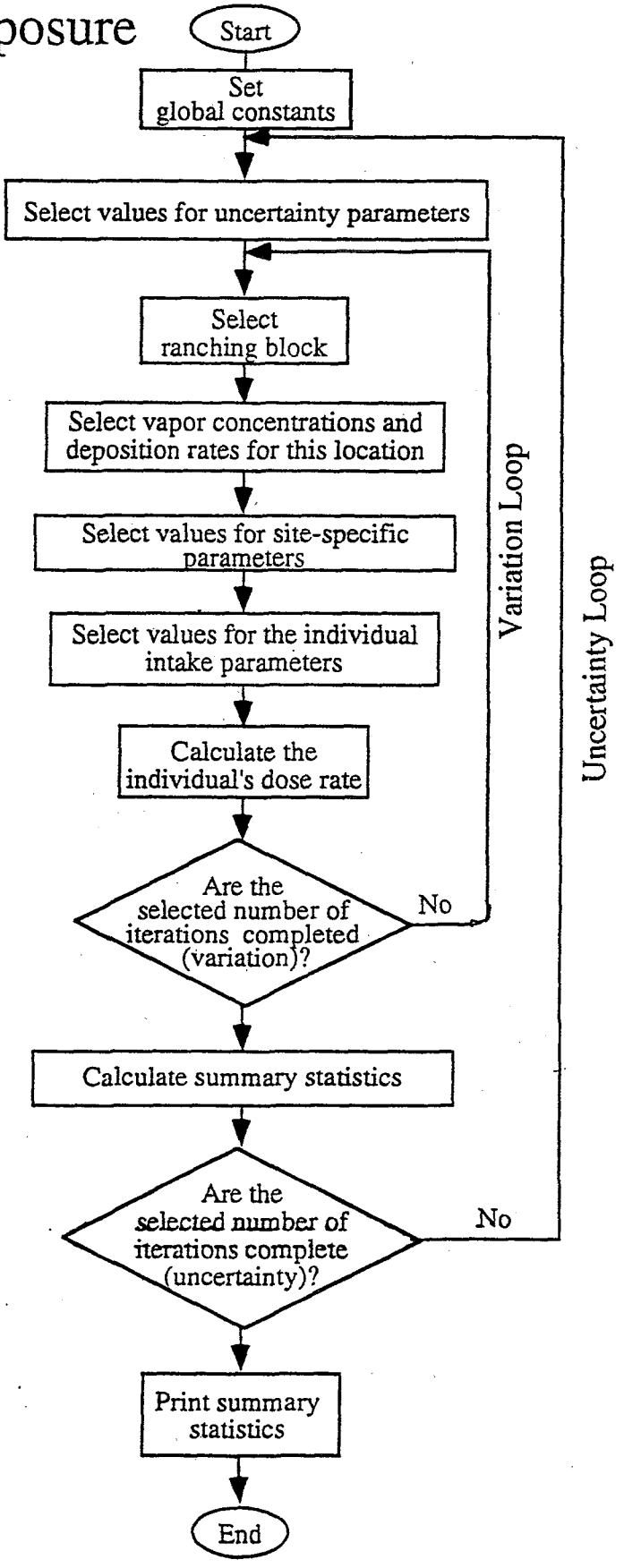
| Uncertainty | Variability |
|---|---|
| Percent particulate-bound dioxin | Location of ranch |
| Air modeling uncertainty | Location of ranch |
| Fraction of time incinerator operational | Location of ranch |
| Feed lot depuration rate | Location of ranch |
| TCDD vapor-to-grass transfer factor | Location of ranch |
| TCDD grass-to-beef transfer factor | Feed lot time |
| Soil loss constant | Beef consumption rate |
| Grass surface weatherization rate | Fraction of homegrown beef |
| Beef cattle consumption rate of grass | Duration of exposure |

# Two Dimensional Monte Carlo Analysis

## Approach: Nested loop simulation



# Nested-Loop Model of TCDD Beef Exposure



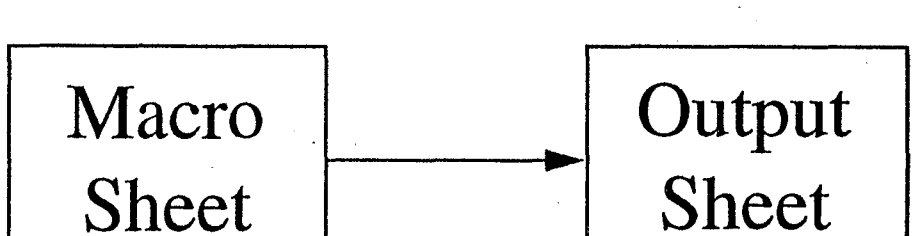
# Software Used to Perform the Analysis

- Microsoft Excel Spreadsheet with Business Graphic and Database: Version 4.0, Mircrosoft Corporation, 1992

- @Risk: Risk Analysis and Simulation Add-In for Microsoft Excel: Version 1.12, Palisade Corporation, 1994

# Model Description

- Written as a series of Excel spreadsheets
- Contains two linked spreadsheets:

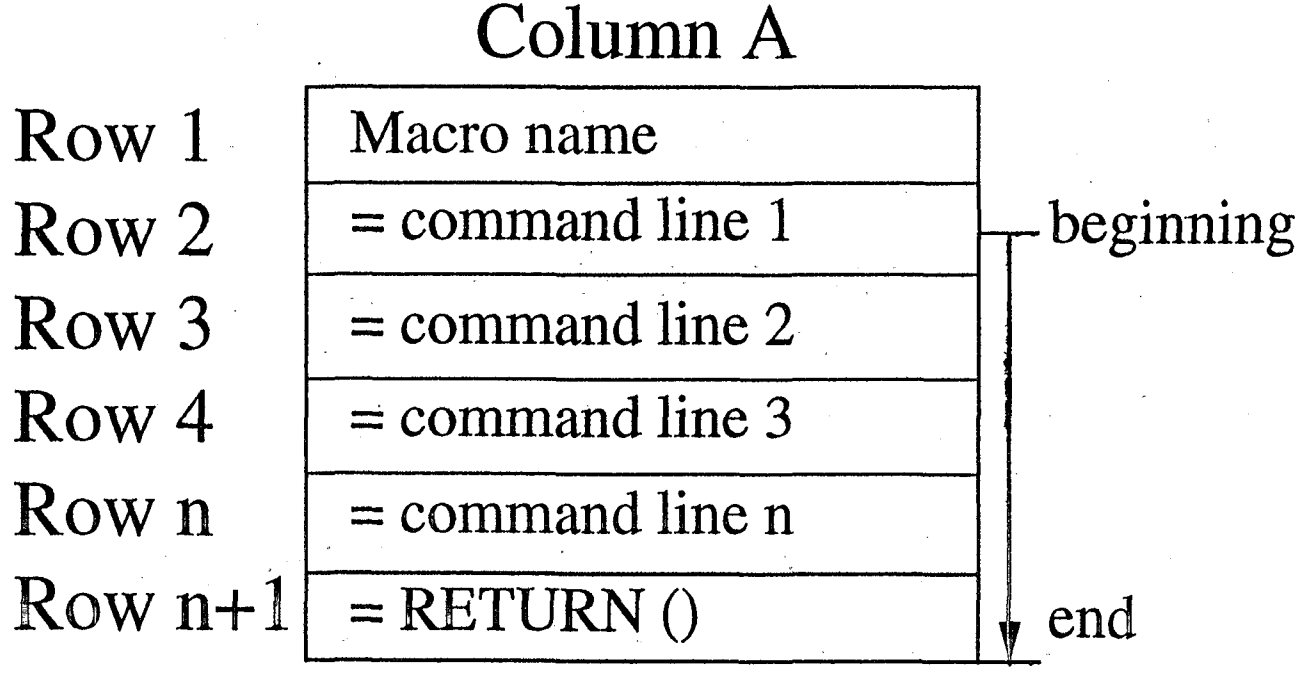


# Macro Sheet

- In a macro sheet, specialized calculations are performed by a series of commands (macros)
- Many macros can be created on one macro sheet
- One macro can call one or more additional macros
- Can use the @Risk function for distributions

# Macro Sheet (cont'd)

- Types of macros used to program Microexposure model are command macros:



# Macro Sheet (cont'd)

Macro commands commonly do the following:

- select a value from a data set a variable name equal to the value
- call other macros
- solve an equation and set a variable equal to the solution
- control the location of an active cell in a worksheet
- end a macro

# Output Sheet

- Receives output from simulation
- Designed by modeler
- Summary statistics can be computer on a range of cells in the Output Sheet at the end of the specified iterations
- Takes the place of the summation and reporting components of @Risk

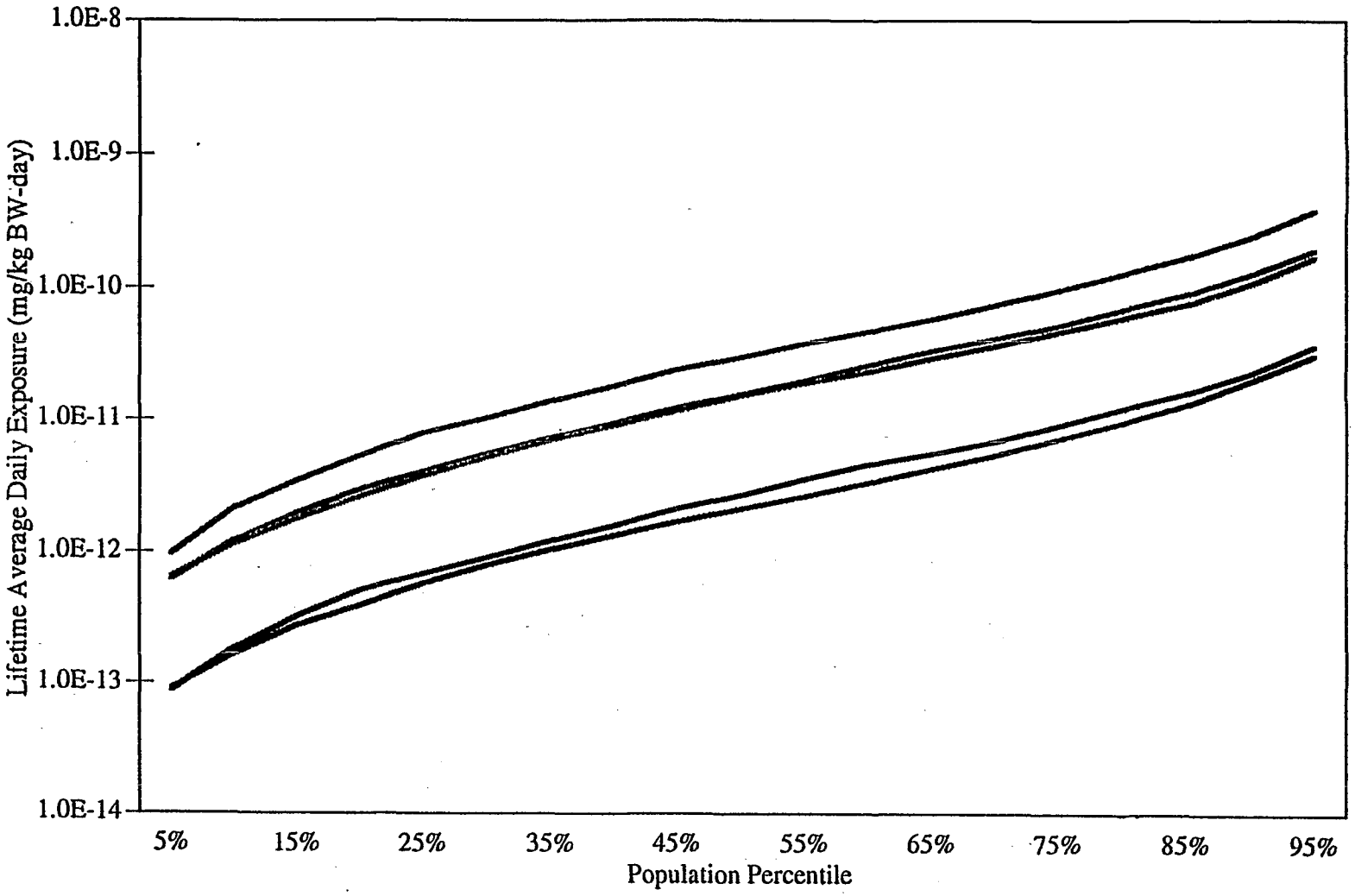
# Data Management

- 2,000 uncertainty loops X 2,000 variation loops = 4,000,000 values
- Only summary statistics (27 values) are and saved for each variation loop
- Summary statistics of the uncertainty in the 2,000 variation loop summary statistics are calculated and saved

# Use of Excel Macros for Two Dimensional Monte Carlo

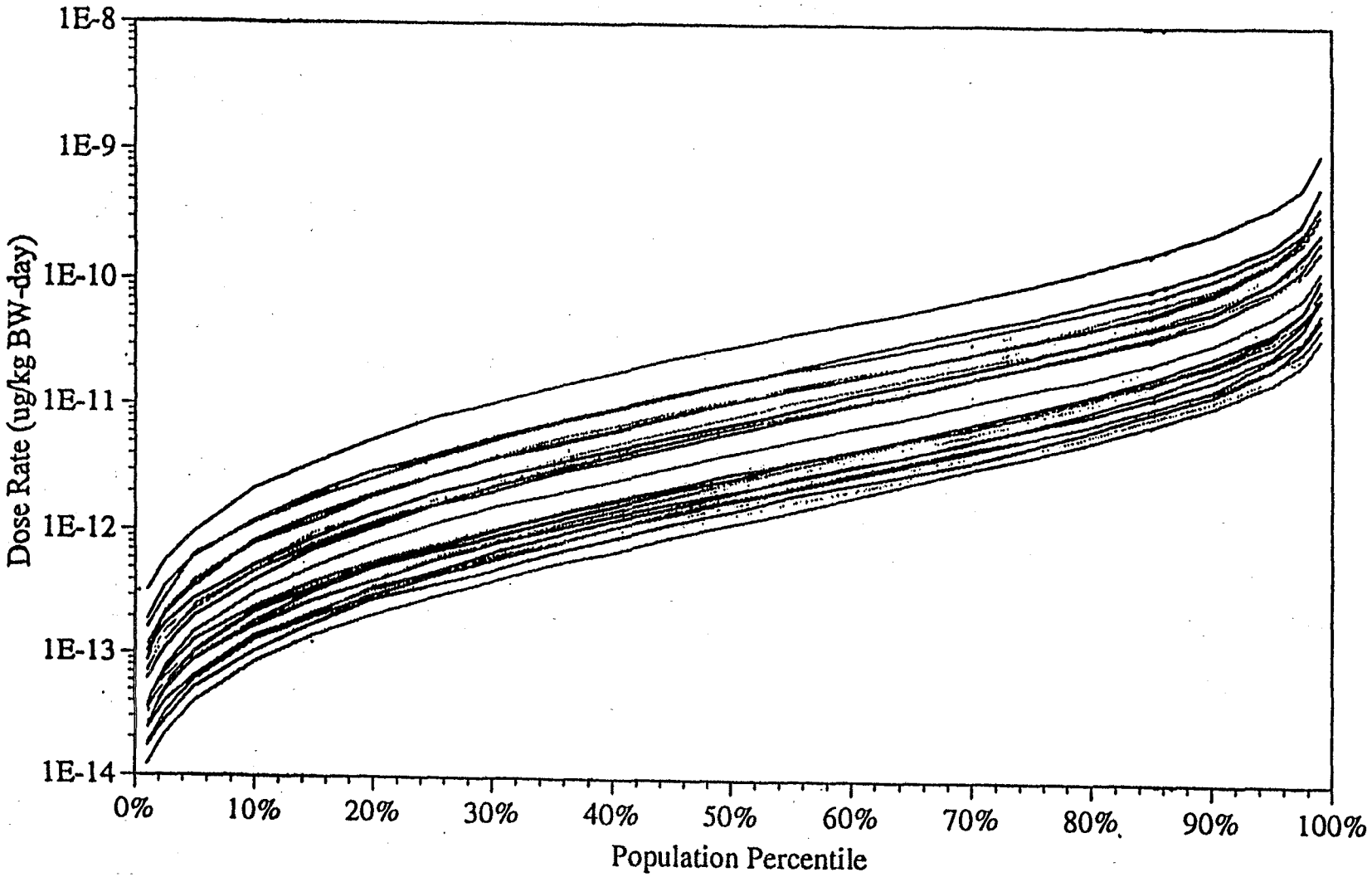
- Advantages
  - Readily available
  - Allows use of @Risk distributions
  - Widely known among risk assessors
  - Good output control
- Disadvantages
  - Hard to de-bug
  - Slow-slow-slow

# RESULTS



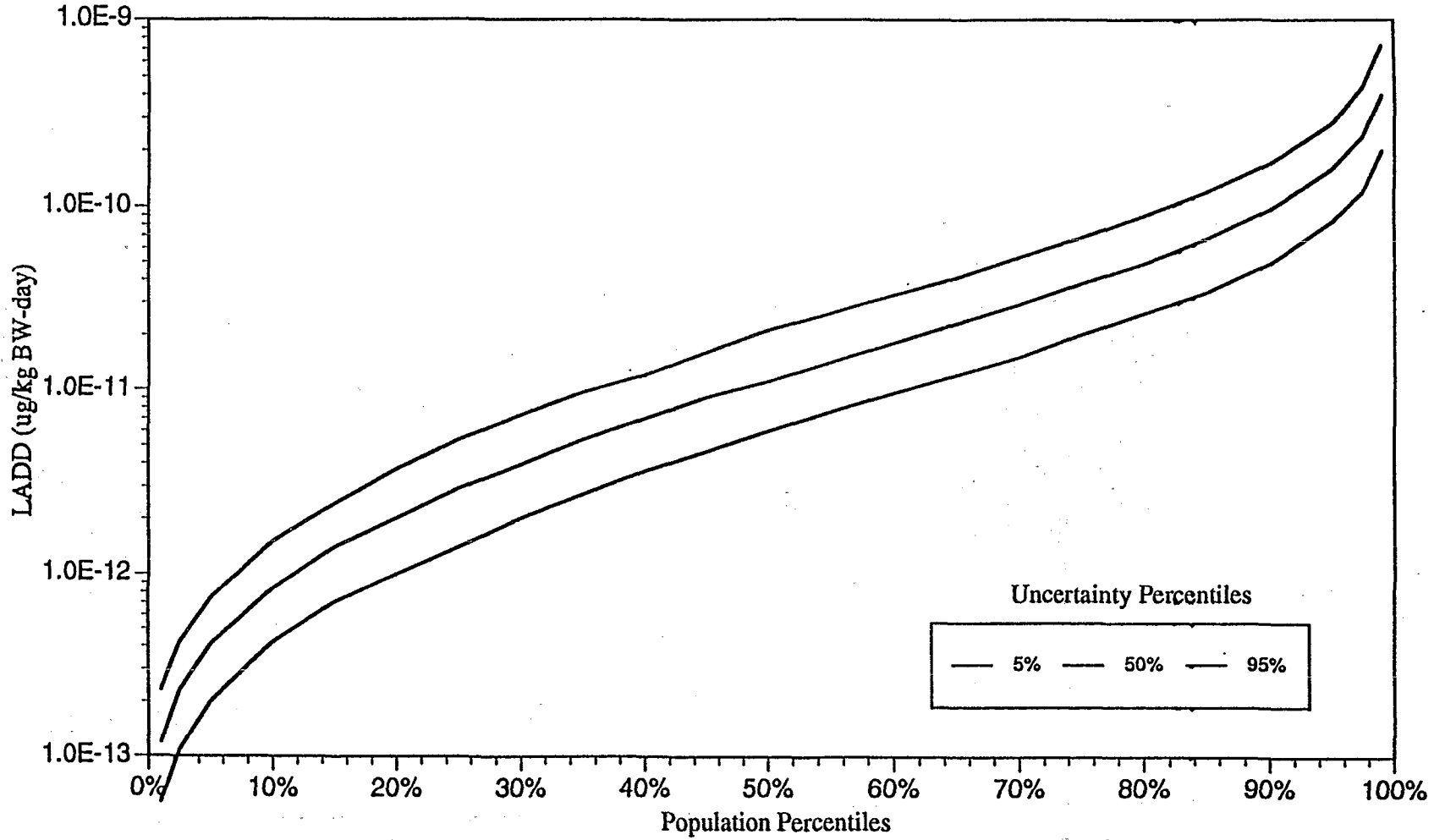

Number of Outer-Loop Iterations: 5

# Output of 25 Uncertainty Loop Runs

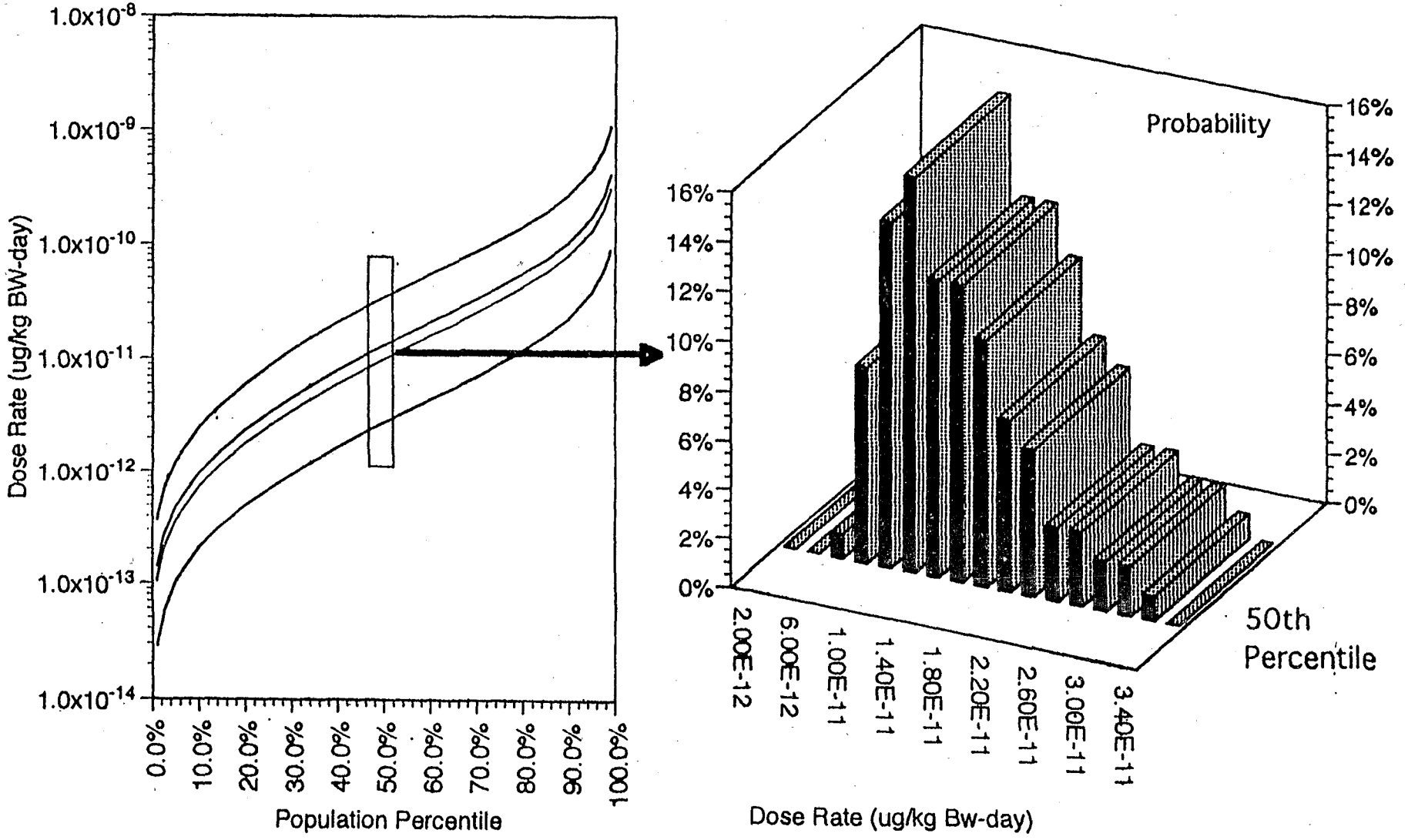



Su-SRA 1994

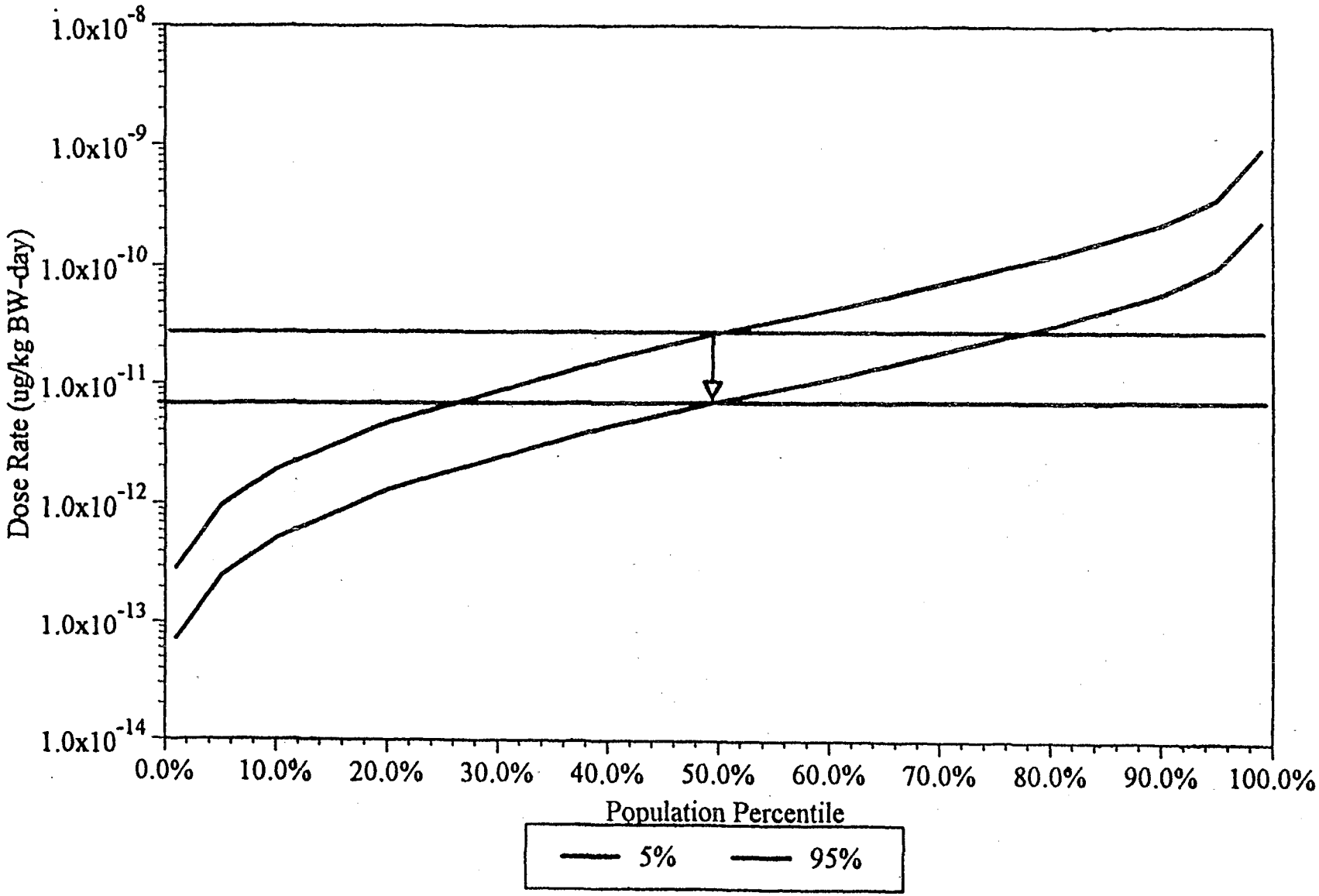
# Variation and Uncertainty in Dose Rates (ug/kg-day) of the Exposed Population for 2,000 Simulations



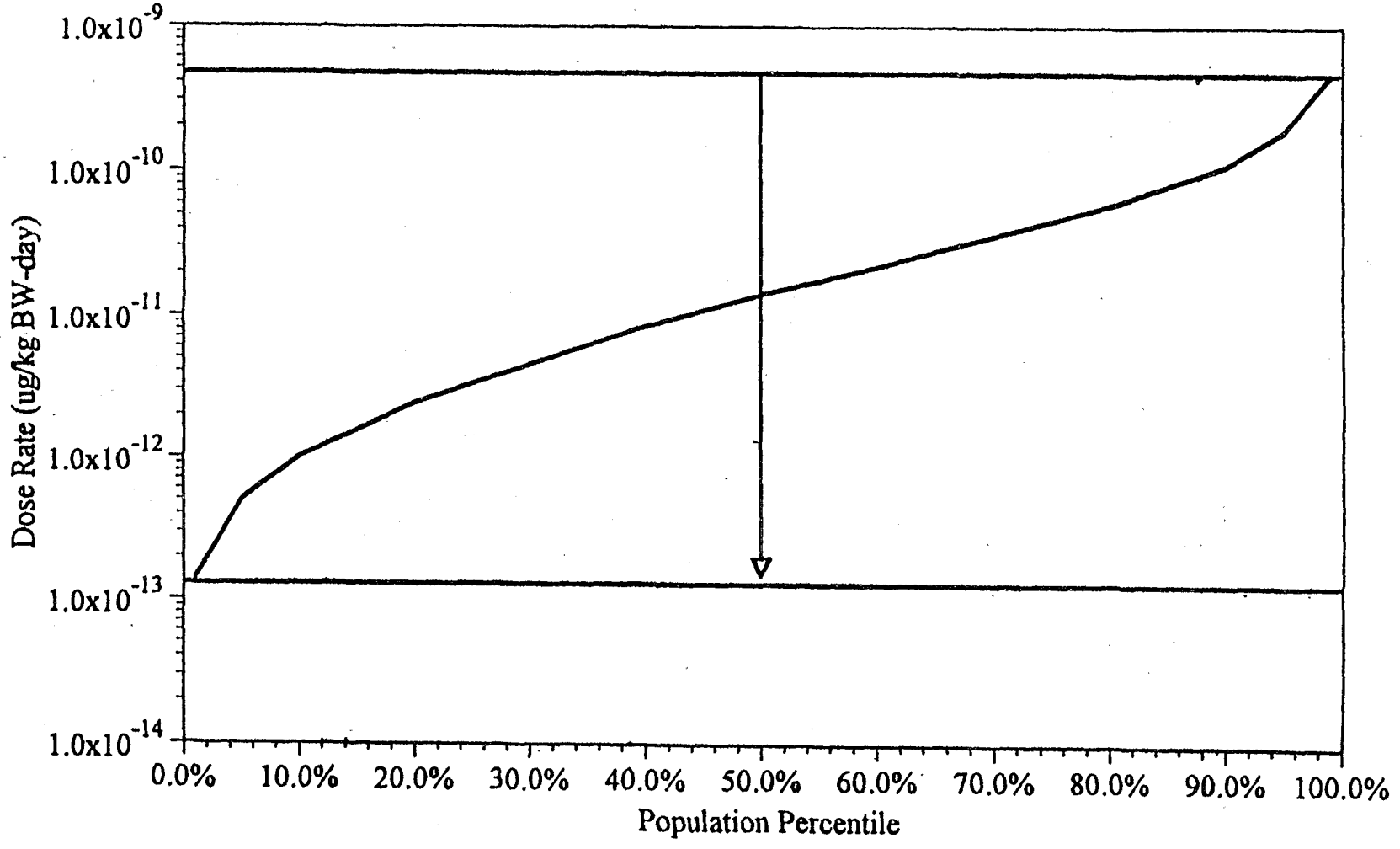
# Results
## An Uncertainty "Slice" of the 50th Percentile Individual



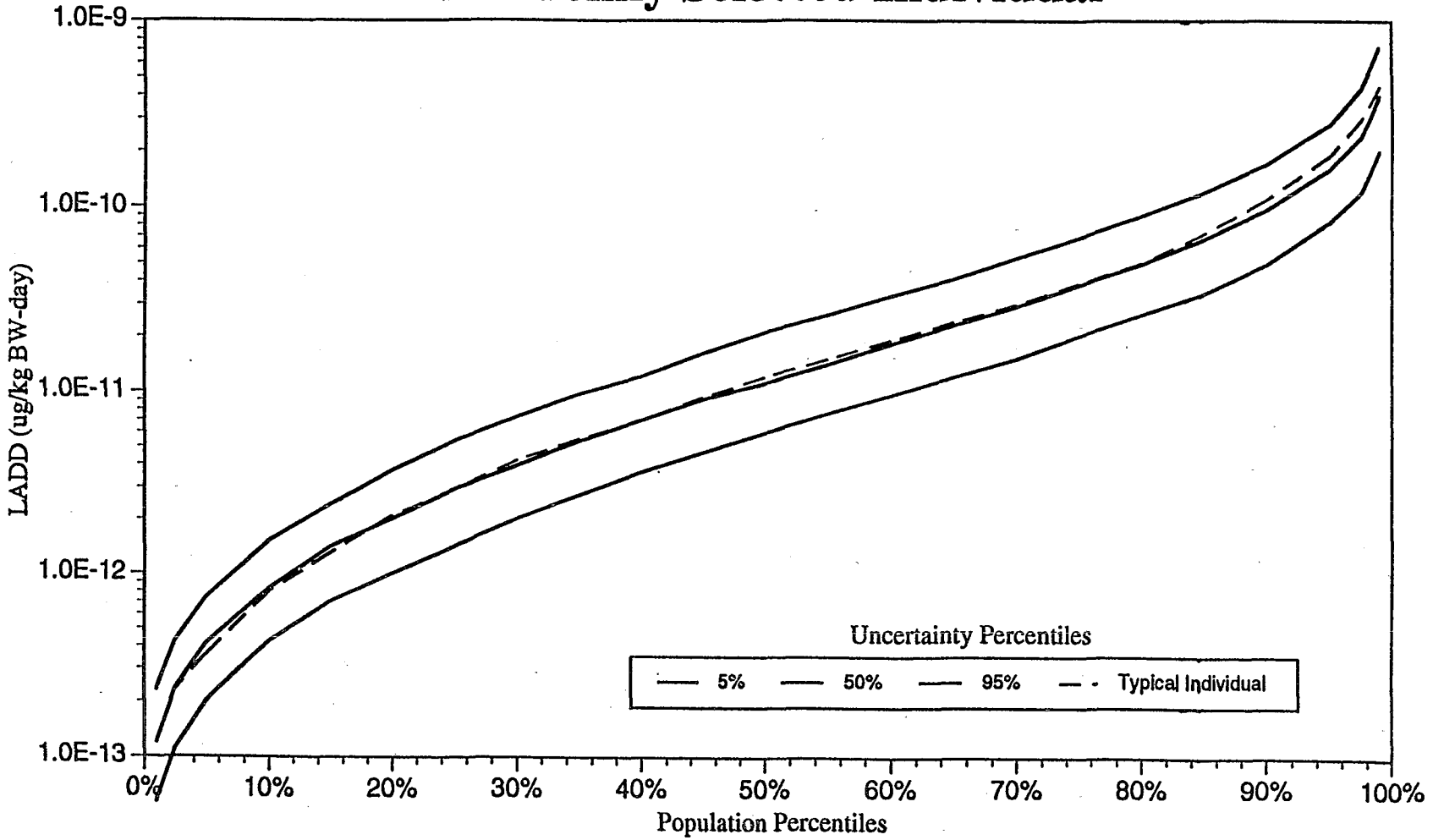
# Results Uncertainty in Exposure Estimates Varies by 1 Order of Magnitude



# Results Interindividual Variation in Exposure Varies by 3 Orders of Magnitude



# Variation and Uncertainty in Dose Rates (ug/kg-day) of the Exposed Population and the Distribution of a Randomly Selected Individual



# The Separation of Parameters into Uncertainty and Variability

- Does not account for the uncertainty component in variability parameters
- Fails to properly evaluate variability/uncertainty interactions
- Future work should take these interactions into consideration through the use of two-dimensional average for variability parameters

# Example: Beef Consumption Rates

- Use the Risk Cumulative function to define a cumulative distribution for beef consumption

= Risk Cumulative (minimum, maximum, ({$\chi_1$,$\chi_2$,…, $\chi_n$,}, {$p_1$, $p_2$, …, $p_n$})

where,

minimum = least possible value

maximum = greatest possible value

$\chi_1$,…,$\chi_n$ = breakpoint values within range

$p_1$,…,$p_n$ = corresponding cumulative probabilities

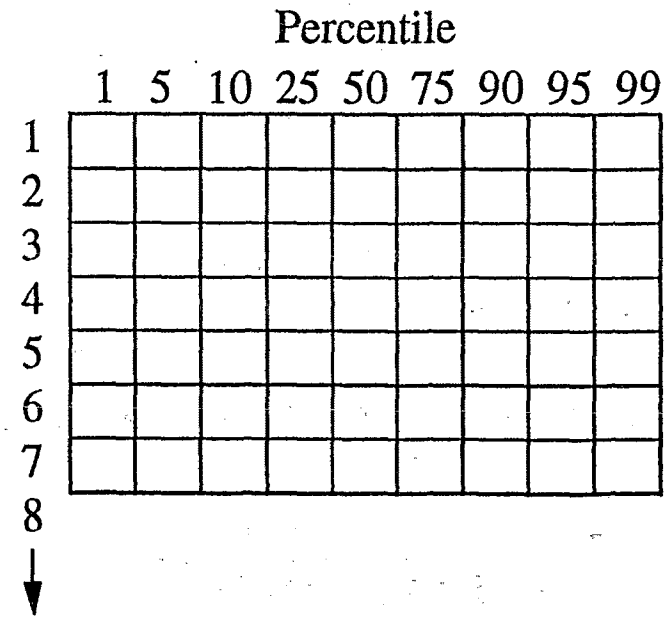
- Used a lookup function in the macro sheet to extract parameter values from a beef consumption uncertainty, variability data array
- Use Horizontal (HLOOKUP) or Vertical (VLOOKUP) function
- In the uncertainty loop define A as the percentile of the uncertainty for beef consumption

- HLOOKUP searches a specified row of an array for a particular value, and returns the value in the indicated cell:

[VALUE] [ARRAY] [ROW]

$p_1$ = HLOOKUP (1, bcrate, A)

Beef Consumption Rate



# Conclusions

- Two-dimensional Monte Carlo analysis can characterize uncertainty and variability in exposure assessments
- Based on the factors considered in this analysis, interindividual variation dominates uncertainty for this exposure pathway
- Because the uncertainty is small there is little difference from the results of a traditional one-dimensional analysis

# CASE STUDY APPLICATION: UNCERTAINTY AND VARIATION IN INDIRECT EXPOSURE ASSESSMENTS: AN ANALYSIS OF EXPOSURE TO TETRACHLORODIBENZO-*P*-DIOXIN FROM A BEEF CONSUMPTION PATHWAY

**Paul S. Price**
**ChemRisk**

First appeared in Risk Analysis 16(2):263-277. Reprinted with permission.

**Case Study :** **Modeling Uncertainty and Variation in Dose Rates from an Indirect Exposure to TCDD from the Consumption of Beef**

This case study will present a Monte Carlo model of indirect exposure to TCDD through the consumption of beef from cattle raised down wind of a hazardous waste incinerator. The results will be presented as a cumulative distribution of individual doses in an exposed population and the uncertainty in that distribution. While this case study involves the use of a large number of parameters and equations, it presents a relatively simple approach for separately evaluating uncertainty and variability in estimates of long-term dose rates.

The case study will demonstrate the following points:

- Use of Excel macro's to allow the use of commonly available software
- An example of a "nested loop" approach for dealing with uncertainty and variability.
- A comparison between the results of the uncertainty and variability model with a monte carlo analysis of total uncertainty.
- Examples of the development of PDFs in the absence of high quality data.

# Uncertainty and Variation in Indirect Exposure Assessments: An Analysis of Exposure to Tetrachlorodibenzo-*p*-Dioxin from a Beef Consumption Pathway

Paul S. Price[1], Steave H. Su[2], Jeff R. Harrington[3], Russell E. Keenan[1]

[1]ChemRisk® - A Division of McLaren/Hart, 1685 Congress Street, Portland, Maine 04102, USA; [2] Bailey Research Associates, 292 Madison Avenue, New York, NY 10017, USA (212) 686-1754; [3] EarthTech, 73 Deering Street, Suite 101, Portland, ME 04101, USA

All correspondence regarding this research should be directed to:

Mr. Paul S. Price
ChemRisk® - A Division of McLaren/Hart
1685 Congress Street
Portland, Maine 04102 USA
(207) 774-0012

## ABSTRACT

Indirect exposures to 2,3,7,8-tetrachlorodibenzo-*p*-dioxin (TCDD) and other toxic materials released in incinerator emissions have been identified as a significant concern for human health. As a result, regulatory agencies and researchers have developed specific approaches for evaluating exposures from indirect pathways. This paper presents a quantitative assessment of the effect of uncertainty and variation in exposure parameters on the resulting estimates of TCDD dose rates received by individuals indirectly exposed to incinerator emissions through the consumption of home-grown beef. The assessment uses a nested Monte Carlo model that separately characterizes uncertainty and variation in dose rate estimates. Uncertainty resulting from limited data on the fate and transport of TCDD are evaluated, and variations in estimated dose rates in the exposed population that result from location-specific parameters and individuals' behaviors are characterized. The analysis indicates that lifetime average daily dose rates for individuals living within 10 kilometers of a hypothetical incinerator range over three orders of magnitude. In contrast, the uncertainty in the dose rate distribution appears to vary by less than one order of magnitude, based on the sources of uncertainty included in this analysis. Current guidance for predicting exposures from indirect exposure pathways was found to overestimate the intakes for typical and high-end individuals.

## KEY WORDS

2,3,7,8-Tetrachlorodibenzo-dioxin, beef, uncertainty, variation, indirect exposure, Monte Carlo

*Risk Analysis*
*An International Journal*

## 1. INTRODUCTION

Indirect exposures to toxic substances released in air emissions can be evaluated by examining those exposure pathways that involve the consumption of animal or vegetable foodstuffs that have accumulated toxic materials via air emissions[1,2]. Previous authors have shown that indirect pathways can result in dose rates that are orders of magnitude higher than the doses received by direct inhalation of vapors and particles[3,4]. Recently, regulatory agencies have taken the position that indirect exposures to toxic substances released in air emissions can represent significant sources of risk to public health. In 1994, EPA issued directives to its regional offices to consider risk from indirect exposure and has since proposed a requirement that indirect risk assessments be performed as part of obtaining new air permits and maintaining existing ones for all incinerators[5]. Three guidance documents for the performance of indirect risk assessments have been released[1,6,7] and information on assessing indirect exposures *to* TCDD is also contained in documents developed as part of the reevaluation of dioxin-like compounds[8].

In these documents, EPA has recommended a series of equations and default parameters with the intent of providing a method for assessing "reasonable" estimates of the dose rates received by a family living at a ranch, given a certain TCDD deposition rate and vapor concentration[1,6,7].

A number of researchers have shown that the use of multiple conservative values or even a blend of "conservative" and "typical" values in an exposure assessment can result in an unknown degree of overestimation of actual risks[9,10]. The potential for overestimating exposure tends to increase for evaluations that involve many parameters or where the parameters are associated with large degrees of uncertainty[11]. Both conditions occur in indirect risk assessments. For example, the calculation of dose rates from beef and dairy consumption may involve more than 40 parameters[1] while only 6 parameters (body weight, air concentration, duration, averaging time, inhalation rate, and lung clearance) are involved in estimating exposures via direct inhalation[12]. EPA[1,6,13] has also acknowledged that many indirect exposure parameters are associated with considerable uncertainty.

When evaluating the imprecision in estimates of exposure, it is important to separately characterize the imprecision that results from interindividual variation and the imprecision that results from

uncertainty in fate and transport processes[10,11,14,15,16,17]. Interindividual variation is the difference in exposure that occurs between one person and another. It can be characterized as the distribution of dose rates in an exposed population. In contrast, uncertainty occurs due to a lack of complete information on a parameter's value. It can be expressed as confidence limits on the dose rate distribution. This characterization does not apply in estimates of individual's dose rates where interindividual variation is simply another source of uncertainty in the estimate of the individual's dose rate[10].

## 2. DESCRIPTION OF THE APPROACH

This paper presents an analysis of the uncertainty in estimates of lifetime average daily dose rates (hereafter called dose rates) potentially received via consumption of beef containing 2,3,7,8-tetrachlorodibenzo-*p*-dioxin (TCDD) emitted from a hypothetical incinerator. For illustrative purposes, this analysis only considers TCDD, and seeks to characterize the uncertainty in the TCDD dose rate estimates developed using EPA methodology[1,6,7]. Since many of the assumptions and parameters are specific to TCDD, the results are not intended to be applied to chlorinated furans or to other dioxin congeners. The beef consumption pathway was chosen because it is the indirect exposure route that produced the highest TCDD dose rates in the EPA's assessment of the WTI incinerator[18]. This analysis does not consider exposure to TCDD via beef that enters the general food supply. Because the processes of cattle transportation, slaughtering, sale, and commercial distribution mix contaminated with uncontaminated beef, such individual exposures are believed to be negligible in comparison to exposures received by ranchers who consume home-raised beef. A nested Monte Carlo analysis is used to quantify the uncertainties associated with the models and to characterize the distribution of dose rates in the exposed population. While this analysis evaluates potential human exposure from a hypothetical incinerator located in Texas, it is not intended to present conclusions about any specific location; rather, it is intended to examine the uncertainty associated with dose rates that are estimated using current guidance.

## 2.1. Description of the Emission Source and the Exposed Population

The hypothetical incinerator is assumed to be located at a point that is equidistant from the cities of Victoria and Freeport on the Gulf Coast of Texas at 28.87° north latitude and 96.19° west longitude. This location was selected because of the availability of meteorological and topographic data and the petrochemical industry present in the region. The incinerator is assumed to emit TCDD as airborne vapor and contaminated particulates at an average rate of $7.6 \times 10^{-10}$ g/sec. This rate is believed to be reasonable for a large incinerator.

In this analysis, the COMPDEP air dispersion model[1,6] is used to characterize transport of TCDD from the stack to the local pastures. The total deposition rates and vapor concentrations at a series of receptor points over a 20 km by 20 km area around the facility were calculated through the use of this model.

The analysis uses the estimated vapor concentration and rates of wet and dry deposition of TCDD to predict concentrations in grass and soil at local pasturage. These concentrations are then used to estimate the TCDD concentration in the edible portion of beef cattle grazed on those pastures. The biotransfer of TCDD from grass and soil to beef was performed using the methodology proposed in the indirect exposure guidance[1,6,7]; however, as discussed below, a TCDD-specific value is used for the grass to beef biotransfer factor.

Cattle ranching occurs throughout this region of Texas (Engbrock, Personal Communication; Lesiker, Personal Communication). Because the Gulf Coast climate is capable of supporting pasture throughout the year, we assume that cattle are raised on pasture exclusively until just prior to slaughter. At that time, cattle are moved to a "feed lot" and are fed a diet of grain in order to increase their weight and the quality of their beef[3].

## 2.2. Evaluation of Model Parameters

In this analysis, the parameters used in the indirect exposure models were separated into three categories: point estimates, "uncertainties", and "interindividual variabilities". The first category contains those parameters that can be anticipated to be known (with a low degree of uncertainty) by the operator of the facility. These values are treated as point estimates in the model. Examples of

such parameters include stack height, stack temperature, average TCDD emission rates, and local topography.

The second category of parameters contains those factors for which there is limited knowledge but for which there is a single true value. Examples of these parameters are the half-life of TCDD in local soils, air-grass partition factors, photodegradation rates, and the biotransfer factors for TCDD from soil and grass to beef. These parameters are modeled in the uncertainty portion of the Monte Carlo model.

The third category includes those parameters that vary from one individual to another within an exposed population. Such variables include the amount of beef consumed by an individual, the fraction of consumed beef raised on contaminated pasture, the location of the ranch relative to the facility, and the duration of time that an individual consumes contaminated beef. These parameters are modeled in the interindividual variability portion of the Monte Carlo model.

Animal-to-animal variation of TCDD in beef (levels in each animal slaughtered) and year-to-year variations in meteorological conditions are not considered in this paper. Because the goal of this analysis is to characterize the long-term variation in doses, such short-term variations are assumed to be effectively averaged out during the course of an individual's long-term exposure. As discussed by Morgan and Henrion[14], estimates of interindividual variability are also subject to uncertainty. This type of uncertainty is not included in this analysis.

A sensitivity analysis was conducted to determine which parameters significantly contributed to the uncertainty and variation in the dose rate estimates. In the sensitivity analysis, each model parameter was increased by 20% from its recommended value and the difference in the estimated dose rate, resulting from the increase of the parameter, was observed. Parameters that an incinerator operator could not be expected to know, and which varied the dose rate by more than 1%, were included in the Monte Carlo model as distributions. Parameters that varied the dose rate by less than 1% were treated as point estimates and were assigned the recommended value or a value equal to the median value of the parameter distribution.

# 3. SOURCES OF INFORMATION USED IN THE ASSESSMENT

## 3.1. General Approach

This analysis uses the available guidance for assessing indirect exposures[1,6,7]. Where possible, we have used the same abbreviations for the parameter names as those provided in the guidance documents. The analysis also draws on work by a number of researchers[19,20,21,22,23]. Table 1 presents the equations for the beef consumption pathway, as given in the EPA guidance documents, and the equations used in the Monte Carlo analyses. Table 2 describes the parameters used in the equations and reports their assignment into the categories of point estimate constants, uncertainty parameters, and variability parameters.

The equations used in the Monte Carlo models differ from the Agency's guidance in three ways. First, where TCDD-specific information on biotransfer factors were available, we used these data in lieu of the log $K_{ow}$-based approach recommended by the current guidance. This allows our analyses to avoid the additional modeling uncertainty that comes from regression-based approaches[13]. Second, we included factors that are not quantitatively accounted for in EPA's indirect exposure guidance. These include the possibility that the plant may not be in constant operation, that the ranch with the highest potential exposure to emission sources may not raise beef for home consumption, and that cattle may be fed grain prior to slaughter. Third, certain equations have been modified to consider uncertainty in air models and vapor-to-particulate partitioning.

Table 2 describes the distributions used to characterize the uncertainty and variation in the parameters used in this analysis. The table also provides references for the sources of these data. For certain parameters, we developed unique distributions based on site-specific information. These parameters are discussed below.

## 3.2. Site Information

Values for facility-specific parameters such as emission rate, stack height, etc., were adapted from actual hazardous waste incinerators operating in the State of Texas. Table 3 provides a description of the values of the facility parameters used in the analysis. Incinerators do not operate on a continual basis because of maintenance time and interruptions in the supply of wastes slated for

Table 1. Equations Used for Point Estimate and Monte Carlo Models for Indirect Exposure

| Point Estimate Model | Monte Carlo Model |
|---|---|
| Lifetime Average Daily Dose Rate | Lifetime Average Daily Dose |
| LADD = Iab * 365 * ED / ATc | LADD = Iab * 365 * ED / ATc |
| Daily Intake from Consumption of Beef | Daily Intake from Consumption of Beef |
| Iab = Ab * Fab * Cab | Iab = Ab * Fab * Cab |
| Concentration of Beef | Concentration in Beef |
| Ab = (QPgB * Pg * Fg + QsB * Sc) * Bag | Ab = [ (QPgB * Pg * Fg * Bag) + (QsB * Sc * Bas) ] * [1 - exp(-Kfl * FLT)] |
| Soil Ingestion Rate of Cow | Soil Ingestion Rate of Cow |
| QsB = 0.03 * Qpg | QsB = 0.03 * Qpg |
| Biotransfer From Grass and Soil to Beef | Biotransfer from Soil to Beef |
| log Bag (beef) = log Bas(beef) = -7.6 + log Kow | Bas = Bag * Sbio/Gbio |
| Concentration of Grass | Concentration in Grass |
| Pg = Pdg + Pvg | Pg = Pdg + Pvg |
| Concentration in Grass Due to Deposition | Concentration in Grass Due to Deposition |
| Pdg = 1000 * [Dyd + * Dyw] * Rpg * [1 - exp(-kp * Tpg)] * 1/Ypg * 1/kp | Pdg = 1000 * [Dyd + * Dyw] * Rpg * [1 - exp(-kp * Tpg)] * 1/Ypg * 1/kp |
| Degradation Rate on Plant | Degradation Rate on Plant |
| kp = Kweath + Kphdeg+volat | kp = Kweath |
| Intercept Fraction of Grass | Intercept Fraction of Grass |
| Rpg = 1-exp(-2.88*Ypg) | Rpg = 1-exp(-2.88*Ypg) |
| Concentration in Grass Due to Air-to-grass Transfer | Concentration in Grass Due to Air-to-grass Transfer |
| Pvg = Cy * Bvg * 1/pa | Pvg = Cy * Bvg * 1/pa |
| Mass-based Air-to-grass Biotransfer Factor | Mass-based Air-to-grass Biotransfer Factor |
| Bvg = (1.19 g/L) * Bvol * 1/0.3 * 1/(890 g/L) | Bvg = (1.19 g/L) * Bvol * 1/0.3 * 1/(890 g/L) |
| Volumetric Air-to-leaf Biotransfer Factor | |
| log Bvol = 1.065 * log Kow - log(H/R*T) - 1.654 | |
| Bvol (2,3,7,8-TCDD) = Bvol /10 | |
| Concentration in Soil | Concentration in Soil |
| Sc = (Dyd + Dyw + Ldif) * [1 - exp (ks * Tc)] * 100 * 1/Z * 1/BD * 1/ks | Sc = (Dyd + Dyw + Ldif) * [1 - exp (ks * Tc)] * 100 * 1/Z * 1/BD * 1/ks |
| Atmospheric Diffusion Flux to Soil | Atmospheric Diffusion Flux to Soil |
| Ldif = 0.31536 * Kt * Cy | Ldif = 0.31536 * Kt * Cy |
| Gas Phase Mass Transfer Coefficient | Gas Phase Mass Transfer Coefficient |
| Kt = 0.482 * u^0.78 * N^-0.67 * de^-0.11 | Kt = 0.482 * u^0.78 * N^-0.67 * de^-0.11 |
| Schmidt Number for Gas Phase | Schmidt Number for Gas Phase |
| N = ua * 1/pa * 1/Da | N = ua * 1/pa * 1/Da |
| Vapor Concentration | Vapor Concentration |
| Cy = AMCy * FV * POf | Cy = AMCy * FV * POf * AIRUNC |
| Dry Deposition Rate | Dry Deposition Rate |
| Dyd = AM Dyd * (1-FV) * POf | Dyd = AM Dyd * (1-FV) * POf *AIRUNC |
| Wet Deposition Rate | Wet Deposition Rate |
| Dyw = AM Dyw * (1-FV) * POf | Dyw = AM Dyw * (1-FV) * POf * AIRUNC |
| Fraction of Vapor | Fraction of Vapor |
| FV = 1- (C x ST/(Psliqt + c x Psliqt)) | FV = 1- (C x ST/(Psliqt + c x Psliqt)) |
| Solid TCDD Vapor Pressure Antoine Equation | Solid TCDD Vapor Pressure Antoine Equation |
| ln Ps = (A - B)/T | ln P = (A - B)/T |
| Solid to Sub-cooled TCDD Vapor Pressure Conversion | Solid to Sub-cooled TCDD Vapor Pressure Conversion |
| Psliq = Ps x EXP(6.79 ((Tm - T)/T) | Psliq = Ps x EXP (6.79((Tm - T)/T) |

Table 2. Exposure Parameters, Parameter Values, and Sources

| Variables | Parameter Type (a) | Definition | Unit | Point Estimate Model Parameters | Point Estimate Model Source (c) | Monte Carlo Model Parameters (d) | Monte Carlo Model Source |
|---|---|---|---|---|---|---|---|
| Ab | M | Concentration in beef | ug pollutant / g | M | Modeled; IEA Eq 5-19 | M | Modeled; IEA Eq 5-19 |
| AIRUNC | U | Air modeling uncertainty | | NA (b) | *RiskTriang(0.6,1,2) | | Earth Tech. (24) |
| AM Cy | V | Air modeled vapor Concentration | ug / $m^3$ | 1.97E-10 | Modeled; COMPDEP for worst case location | M | Earth Tech. (24) |
| AM Dyd | V | Air modeled dry deposition rate | g pollutant / $m^2$ - yr | 3.03E-11 | Modeled; COMPDEP for worst case location | M | Earth Tech. (24) |
| AM Dyw | V | Air modeled wet deposition rate | g pollutant / $m^2$ - yr | 1.42E-11 | Modeled; COMPDEP for worst case location | M | Earth Tech. (24) |
| ATc | C | Carcinogenic averaging time | day | 25550 | | 25550 | |
| Bag | U | Biotransfer factor for grass | d / kg | 1.10E-01 | Travis and Arms (19); IEG 5-31 | Risk Discrete ({0.055,0.076,0.079},{1,1,1}) | Jensen et al. (23) |
| Bas | U | Biotransfer factor for soil | d / kg | 1.10E-01 | Travis and Arms (19); IEG 5-31 | M | Modeled |
| BD | U | Soil bulk density | g / $cm^3$ | 1.5 | IEG 4-11 | *RiskTriang (0.93,1.5,1.84) | IEG 4-11 |
| Bvg | M | Mass-based air-to-plant biotransfer factor | | M | Modeled; IEA Eq 5-15b | M | Modeled; IEA Eq 5-15b |
| Bvol | M | Volumetric-based air-to-grass biotransfer factor | | M | Modeled; Bacci (35); IEA 5-9 | $10^{6.9}$ | McCrady and Maggard (28); IEA 5-9 |
| c | C | Junge's constant | atm-cm | 1.7E-4 | IEA 3-35 | 1.7E-4 | IEA 3-35 |
| Cab | V | Consumption rate of beef | g DW/kg BW-day | 0.555 | McKone and Ryan (22) | *RiskCumul{0.363,2.73,{1.01,1.12,1.2,1.26,1.32,1.36, 1.41,1.45,1.49,1.53,1.57,1.61,1.65,1.7,1.74,1.8,1.86,1.93,2.05}, {0.05,0.1,0.15,0.2,0.25,0.3,0.35,0.4,0.45,0.5,0.55,0.6,0.65,0.7,0.75,0.8,0.85,0.9,0.95}} | McKone and Ryan (22) |
| Cy | M | Vapor concentration of dioxin | ug / $m^3$ | M | Modeled based on COMPDEP | M | Modeled based on COMPDEP result |
| Da | U | Diffusion coefficient of pollutant in air | $cm^2/s$ | 0.0525 | Estimated based on Lyman et al., 1982 | RiskTriang (0.0485,0.0525, 0.0565) | Estimated based on Lyman et al. (36) |
| de | C | Effective diameter of contaminated area | m | 2856 | Calculated based on 475 acre pasture | 2856 | Lesiker, Personnel Communication |
| Dyd | M | Yearly dry deposition rate | g pollutant / $m^2$ - yr | M | Modeled based on COMPDEP result | M | Modeled based on COMPDEP result |
| Dyw | M | Yearly wet deposition rate | g pollutant / $m^2$ - yr | M | Modeled based on COMPDEP result | M | Modeled based on COMPDEP result |
| ED | V | Exposure duration | yr | 30 | IEG | *RiskCumul(1,70, {2.4,10,26.7,48.3,58.4}, {0.25,0.5,0.75,0.9,0.95}) | Israeli and Nelson (25) |
| Fab | V | Fraction of local/contaminated beef | | 0.44 | USDA (37); IEA 5-15 | RiskTriang (0.01,0.44,1) | USDA (37); IEA 5-15 |
| Fg | C | Fraction of consumed grass contaminated | | 1.0 | Assumed 100% grass consumed comes locally | 1.0 | Assumed 100% |
| FLT | V | Feed lot time | day | NA | | *RiskTriang (30,120,360) | EED 10-90; Jensen et al. (23) |
| FV | U | Fraction of dioxin emission vapor bound | | M | Lorber et al. (2) | M | Lorber et al. (2) |
| Fw | C | Fraction wet deposition adhering to grass | | 1 | IEA 5-5 | 1 | IEA 5-5 |
| Gbio | U | Grass bioavailability | | NA | | *(0.45,0.55) | Fries and Paustenbach (3) |
| H | | Henry's Law Constant | atm-$m^3$/mol | 1.6 | EED A-3 | | |

## Table 2. Exposure Parameters, Parameter Values, and Sources

| Variables | Parameter Type (a) | Definition | Unit | Point Estimate Model | | Monte Carlo Model | |
|---|---|---|---|---|---|---|---|
| | | | | Parameters | Source (c) | Parameters (d) | Source |
| Iab | M | Annual intake from beef | ug pollutant / kg BW - day | M | Modeled; IEG Eq 5-24 | M | Modeled; IEG Eq 5-24 |
| Kfl | U | Feed lot reduction constant | day$^{-1}$ | NA | | *RiskNormal (0.006,0.00043) | EED 10-90; Jensen et al. (23) |
| kp | M | Grass surface loss constant | yr$^{-1}$ | M | Kweath | M | Kweath |
| Kphdeg+ volat | U | Grass surface loss from photodegradation and volatilization | yr$^{-1}$ | 179 | McCraddy and Maggard (28) | RiskNormal (179,14.9) | McCraddy and Maggard (28) |
| ks | U | Soil loss constant | yr$^{-1}$ | 6.30E-02 | ln2/ soil half-life | *LN(2)/RiskUniform (10,12) | ln2/ half-life |
| Kt | M | Gas phase mass transfer coefficient | cm/s | M | Modeled; IEG Eq 4-6 | M | Modeled; IEG Eq 4-6 |
| Kweath | U | Grass surface loss from weathering | yr$^{-1}$ | 25.3 | Baes et al. (27); EED 5-47 | *RiskTriang (7.44,25.3,126) | EED 5-47 |
| LADD | M | Lifetime average daily dose | ug/kgBW-day | M | | M | |
| Ldif | M | Atmospheric diffusion flux to soil | g/m$^2$-yr | M | Modeled; IEA Eq 4-1, p4-2 | M | Modeled; IEA Eq 4-1, p4-2 |
| logKow | U | logOctanol-water partition coefficient of pollutant | | 6.64 | Marple (28); EED 2-8 | | |
| N | M | Schmidt number for gas phase | | M | Modeled; IEG Eq 4-7 | M | Modeled; IEG Eq 4-7 |
| pa | C | Density of air | g/m$^3$ | 1190 | IEG 4-24 | 1190 | |
| Pdg | M | Concentration in grass due to deposition | ug pollutant / g grass DW | M | Modeled; IEG Eq 5-4 | M | Modeled; IEG Eq 5-4 |
| Pg | M | Concentration in grass | ug pollutant / g DW | M | Modeled; IEG Eq 5-1 | M | Modeled; IEG Eq 5-1 |
| POf | U | Fraction of time incinerator is operational | | 1 | Assumed | *RiskTriang(0.6,0.8,1) | Assumed |
| Pvg | M | Concentration in grass due to air-to-plant transfer | ug pollutant / g grass DW | M | Modeled; IEA Eq 5-13 | M | Modeled; IEA Eq 5-13 |
| QPgB | U | Beef cow consumption rate of grass | kg DW / day | 12 | IEG 5-26 | *RiskTriang(6,12,18) | IEG 5-26 |
| QsB | U | Beef cow consumption rate of soil | kg DW / day | 0.36 | IEG 5-28 | M | IEG 5-28 |
| R | C | Ideal Gas Constant | atm-m$^3$/mole-deg K | 8.21E-05 | IEA 5-8 | | |
| Rpg | M | Intercept fraction of edible portion of grass | | M | Modeled; IEG Eq 5-5; from Baes et al. (27) | M | Modeled; IEG Eq 5-5; from Baes et al. (27) |
| Sbio | U | Soil bioavailability | | | NA | *RiskUniform(0.3,0.4) | Fries and Paustenbach (3) |
| Sc | M | Concentration in soil | ug / g | M | Modeled; IEA Eq 4-1 | M | Modeled; IEA Eq 4-1 |
| $S_T$ | U | Aerosal surface area | cm$^2$/cm$^3$ | 3.5E-6 | IEA 3-35; Lorber et al. (2) | Risk Triang (6.5E-12, 3.0E-11, 4.3E-11) | IEA 3-35; Lorber et al. (2) |
| T | C | Temperature | deg K | 298.1 | IEA 5-8 | Risk Cumul(261,312, {282,287,290,293,296,298,298,301,304}, {0.1,0.2,0.3,0.4,0.5,0.6,0.7,0.8,0.9} | Local meteorological data |
| t(1/2) | U | Soil Half-life of pollutant | yr | 11 | Mackay et al. (39) | *RiskUniform(10,12) | Mackay et al. (39) |
| Tc | U | Total deposition time to soil | yr | 80 | Plant operation life | *RiskTriang(10,80,100) | Plant operation life |
| Tm | C | Melting temperature | deg K | 578 | Schroy et al. (40) | 578 | Schroy et al. (40) |
| Tpg | C | Grass exposure time to deposition per harvest | yr | 0.123 | IEG 5-13 | 0.123 | IEG 5-13 |
| u | U | Wind speed | m/s | 4.61 | EarthTech (24) | *RiskCumul(0,17.49, {2.06,2.57,3.09,3.6,4.12,4.63,5.66, 6.69,7.72},{0.1,0.2,0.3,0.4,0.5,0.6,0.7,0.8,0.9}) | EarthTech (24) |
| ua | C | Viscosity of air | g/cm-s | 184 | IEG 4-24 | 184 | IEG 4-24 |
| Ypg | U | Yield (crop biomass) | kg DW/m^2 | 0.31 | IEG 5-15; Belcher and Travis (20) | *RiskTriang(0.02,0.31, 0.75) | IEG 5-15; Belcher and Travis (20) |
| Z | U | Soil mixing depth | cm | 1 | Assumed | *RiskTriang(0.5,1,1.5) | Assumed |

a. Variable types: C, constant; M, modeled; U, uncertainty; V, variability
b. NA, Not applicable
c. Sources: EED, Estimating Exposure to Dioxin-like Compounds (8); IEA, Indirect Exposure Addendum (6); IEG, Indirect Exposure Guidance (1)
d. Distributions are characterized using the parameter formats from @Risk (38)

**Table 3. Specifications Used for the Hypothetical Facility (a)**

| Parameter | Value | Units |
|---|---|---|
| Stack Height | 30 | m |
| Base Elevation | 9 | m |
| Stack Diameter | 0.76 | m |
| Stack Temperature | 425 | K |
| Exhaust Flow Rate | 6.91 | acms (b) |
| Exit Velocity | 15.24 | m/s |
| Building Height | 6.1 | m |
| Building Width | 6.1 | m |
| Building Length | 15.24 | m |
| Total TCDD Emission Rate | 0.76 | ng/sec |

a. Values are based on typical incinerators in the state of Texas.

b. Actual cubic meters per second.

burning; as a result, the assumption that emissions occur on a constant basis is not appropriate. In this analysis we have assumed, based on discussions with incinerator operators in Texas, that the plant will be in operation for at least 60% of the time. The fraction of time the plant is in operation is given by a triangular distribution with a range of 0.6 to 1.0, with a most likely value of 0.8.

Topographic information for the vicinity of the hypothetical incinerator was based on models produced by the United States Geological Survey. This analysis used meteorological data from the National Weather Service Station in Victoria, Texas, for the 5-year period of 1985 to 1989.

Information on ranch sizes and practices was obtained by contacting the county extension agents in the two counties (Matagorda and Jackson) in the area of hypothetical impact (Lesiker, Personal Communication; Engbrock, Personal Communication). Table 4 presents the information obtained concerning local ranching activities.

In estimating the dose rates received by the potentially exposed population, it must be considered that most ranches do not slaughter cattle for home consumption. Based upon information received from Federal, State, and local experts, the number of ranches in the U.S. that slaughter their own beef for home consumption is very small (Simpson, Personal Communication; Brown, Personal Communication[3,4]). In addition, county extension agents have indicated that the majority of ranches in the area of the facility raise calves not beef animals (Lesiker, Personal Communication; Engbrock, Personal Communication). Based upon this information, we have made the assumption that one ranch in twenty (5%) in the area will slaughter of cattle for home consumption.

The duration of an individual's exposure was assumed to equal the period of time that an individual lives on an affected ranch. Information on the distribution of such residence times is taken from national estimates of residence times for individuals living on farms and ranchers[25].

### 3.3. Air Dispersion Modeling

The COMPDEP model has been shown to produce predicted air concentrations that compare favorably to observed levels. The confidence in the model predictions is greater for estimates of long-term concentrations (yearly or longer) than for short-term estimates (hourly or daily). Estimates by EPA (40 CFR part 51 Appendix W) indicate that errors of ±40% are found in

**Table 4. Cattle Ranching Practices in Matagorda and Jackson County, Texas**

| | |
|---|---|
| • | 35,000 mature cows in Jackson County |
| • | 752 ranches |
| • | Average ranch size is 475 acres |
| • | Average stocking rate is 5 acres per head of cattle |
| • | All ranches only use pasture or some type of forage |
| • | All ranches are considered cow-calf operations, or "breeding herds." |
| • | Very few (<5%) of the ranchers raised beef for home consumption |

estimates of the highest predicted concentrations. A new variable, AIRUNC was created to account for modeling uncertainty. This variable is defined by a triangular distribution with a most likely value of 1.0, a minimum of 0.6, and a maximum of 1.4.

When determining the vapor/particulate partitioning of TCDD, modeling is preferred over monitoring data because of uncertainties in the sample collection methodology[6,8]. A theoretical method has been suggested by Bidleman[26] for the TCDD fraction which is not permanently bound to particulates and freely exchanges between the vapor and particulate-bound form. The Bidleman approach involves the estimation of particulate-bound fraction based on the physical chemical properties of the compound, the surface area of particles in the air and the ambient temperatures.

In order to reflect the particle surface area of airborne particulates in a rural farming area, we applied a triangular distribution representing "clean background," "average background," and "background plus local sources" of $4.2 \times 10^{-7}$, $1.5 \times 10^{-6}$, and $3.5 \times 10^{-6}$ $cm^2/cm^3$, respectively as described in Bidleman[26]. Using this approach, we estimated that the fraction of TCDD that remains in the vapor form ranges widely based on the particle surface area and the endpoint temperature. At the average local temperature the fraction that remains in a vapor form is 91% for the low particulate "clean background" and 53% for the higher particulate "background plus local sources" scenarios.

In this analysis, the COMPDEP model was run twice. The first run assumed that the TCDD was released as a vapor, while the second modeled TCDD releases in a bound form. The vapor concentration predicted for a given location is calculated by taking the COMPDEP air modeling results and multiplying them by AIRUNC, the fraction of TCDD released as a vapor, and the fraction of the time the facility is in operation. A similar approach is used to characterize the uncertainty in particulate bound TCDD (both wet and dry) deposition rates at a specific location.

### 3.4. Plant and Animal Uptake

The uncertainty in the accumulation of TCDD on plant surfaces is estimated based on the range of weathering rates reported by Baes et al.[27]. The estimated uncertainty in the vapor-to-grass transfer of TCDD is based on the experimental uncertainty in the partitioning of TCDD reported by McCrady and Magard[28]. Distributions reflecting uncertainties for the rates of consumption of soil

and grass by cattle were taken from EPA[1]. Data on the uncertainty in the annual average wind speed was developed from the available meteorological data. The distributions developed for these parameters and the bases for their derivations are provided in Tables 1 and 2.

In this analysis, the biotransfer factor (Bag) used to predict the concentration in beef as a result of dietary uptake by cattle is taken from the work by Jensen et al.[23]. This study is the basis for the TCDD data in Travis and Arms' log $K_{ow}$ regression equation[19] suggested by the EPA guidance[1,6,7]. Jensen et al. reported nondetectable levels of TCDD in muscle tissues in three cows following 28 days feeding of a diet containing 24 ppt TCDD. Although TCDD was not detected in the muscle tissues, it was found to be present in fat tissue. We estimated the TCDD concentration in muscle tissue based on the reported average fat content of the muscle samples and their respective fat-based biotransfer factors for each of the three cows. The value of Bag estimated from original data are 0.055, 0.076, and 0.079 are slightly lower than the value of 0.11 estimated for TCDD using Travis and Arm. The uncertainty in Bag is characterized by discrete distribution of the three values.

Grain has a much lower potential for TCDD contamination from air emissions than pasture[2,3]; as a result, TCDD intake will greatly decline during grain feeding. The following equation is used to predict the change in TCDD concentration in beef that occurs during the time that the animal is fed grain:

$$Ab_{slaughter} = (Ab_{before\ feed\ lot})e^{-(K_{fl}*FLT)}$$

where; $Ab_{slaughter}$ is the concentration of TCDD in beef at slaughter, $Ab_{before\ feed\ lot}$ is the concentration in beef before the cattle are placed on feed lot, $K_{fl}$ is the TCDD depuration constant, and FLT is the length of time the cattle spend on the feed lot. Jensen et al.[23] studied the elimination of TCDD from adipose tissue of beef cattle during the contaminant-free feeding period prior to slaughter. The concentration of TCDD in cattle adipose tissue was found to decrease rapidly, with a half–life of 16.5 ± 1.4 weeks, or a first order rate constant of 0.042 ± 0.003 $week^{-1}$. Since the Jensen et al. elimination rate does not account for diluting the tissue TCDD concentrations accompanying a weight increase, this analysis does not consider the impact of dilution and therefore, overestimates potential exposures.

The USDA Current Agricultural statistics on the number of animals in feed lots can be used to derive an estimate of 4-6 months for the average period cattle are placed on feed lots[29]. For this analysis, we have assumed that the duration of time for grain feeding can be characterized by a triangular distribution, with a most likely value of four months, a minimum value of one month, and a maximum value of six months. We have shortened this range time to reflect the possibility that beef animals raised for home consumption may not be kept "on feed lot" as long as animals raised for commercial slaughter.

## 4. ANALYSES OF UNCERTAINTY AND VARIATION

Three analyses were performed for this paper. First, we developed point estimates of the dose rates for the typical and high-end rancher. Second, we conducted a Monte Carlo model of uncertainty and variation in dose rates for the population of individuals living on ranches in the modeled area. This analysis used a nested Monte Carlo model. Third, we modeled the uncertainty in the dose rates for the individual living on a ranch that receives the highest level of exposure to the incinerator emissions (of the ranches that raise beef for home consumption) and uncertainty in the dose rates for an individual on a ranch receiving the "average" exposure to TCDD emissions. Both analyses used a second Monte Carlo model that did not distinguish between interindividuality and uncertainty.

### 4.1. Development of a Point Estimate of Dose Rates for the High End and Typical Individuals

A point estimate-based analysis of the doses received from the beef consumption pathway is conducted using methodologies recommended in EPA guidance[1,6,7]. The receptor location with the highest estimated TCDD concentration in beef is used in the estimate of the "high-end" individual. The beef concentration for the "typical" individual was the median beef concentration for all the modeled receptor locations. These dose rates represent intakes for a rancher who raises cattle entirely on pasture and are based on the predicted values of vapor-to-plant and plant-to-beef biotransfer provided in EPA's guidance[1,6,7].

## 4.2. Monte Carlo Model of Uncertainty and Variation in the Local Population

A two-dimensional Monte Carlo model that separately characterizes uncertainty and variation was constructed using the nested loop approach described by Hoffman and Hammonds[16] and Barry[30]. The approach uses an iterative procedure (the uncertainty loop) to select values for the uncertainty parameters. During each iteration of the uncertainty loop, the model uses a second nested iterative procedure (the variation loop) which characterizes the human variation in beef consumption in the exposed population. Figure 1 presents a flow chart for the model.

The model begins by randomly selecting a set of values from the probability density functions defined for each of the parameters in the uncertainty loop (see Table 2). Once these parameter values are selected, the model enters the variation loop. The variation loop models a distribution of dose rates in the exposed population determined by the selected set of values for the uncertainty parameters. In each iteration of the variation loop, the model randomly extracts values from the probability density functions that describe parameters that vary among individuals. The variation loop also selects a location where the modeled individual lives and uses the location to determine the appropriate long-term deposition rates and airborne concentrations. The locations have been identified in the following manner:

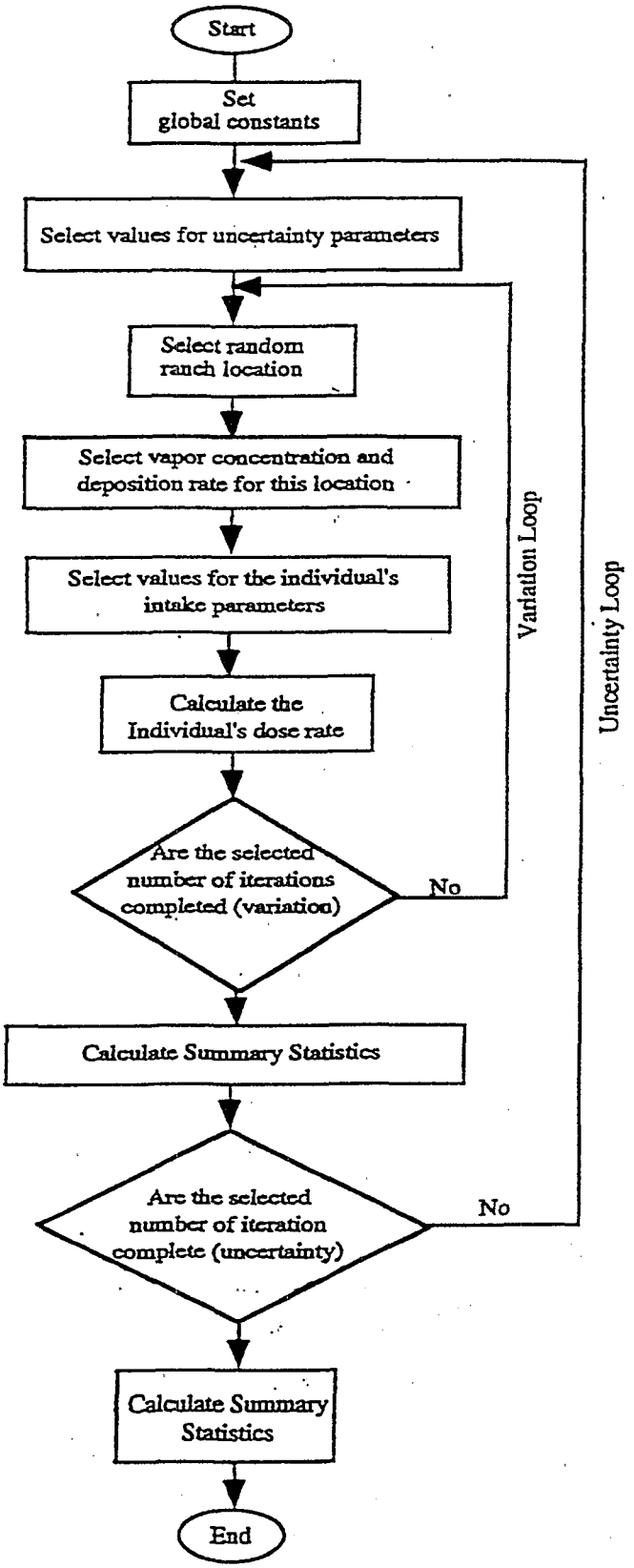
1. The 400 $km^2$ area has been divided into 450 acre blocks of land (the average size of ranches in this region of Texas).

2. Half the blocks of land have been randomly selected to serve as cattle ranching operations; the remainder are assumed to be used for nonranching land uses such as roads, farms, and residential or commercial land uses.

3. The output of the air dispersion model is used to calculate the average airborne concentrations and deposition rates that occur across each of the 450 acre blocks of land.

For each of the modeled individuals, the variation loop selects a set of vapor concentrations and deposition rates for one of the cattle ranching blocks. The model then estimates the dose rate that the individual receives from consuming TCDD-contaminated beef. The variation loop is repeated until the dose rates for the specified number of individuals have been obtained. The model then calculates the summary statistics for the modeled population and stores them for the final output. Once completed, the model returns to the uncertainty loop and selects a new set of values for the

Figure 1. Monte Carlo Model of Uncertainty and Variation of Indirect Exposure in the Local Population



uncertainty parameters. This process is repeated for a specified number of iterations. Latin Hypercube sampling method is used in the model to provide efficient parameter sampling.

A total of 2,000 model iterations were conducted for both the uncertainty and variation loops for the model. Since each uncertainty loop required 2,000 iterations of the variation loop, the model performed 4,000,000 separate dose rate estimates. In separate analyses, we found that 2,000 iterations for each of the loops are sufficient to produce stable estimates for the 90th percent confidence limits of the median, mean, and 95th percentile outputs of the dose rate distribution.

The Monte Carlo model was constructed on the PC-platform using Excel© 4.0 macro language. The Monte Carlo sampling of parameters from probability density functions was accomplished using @Risk Version© 1.123[1].

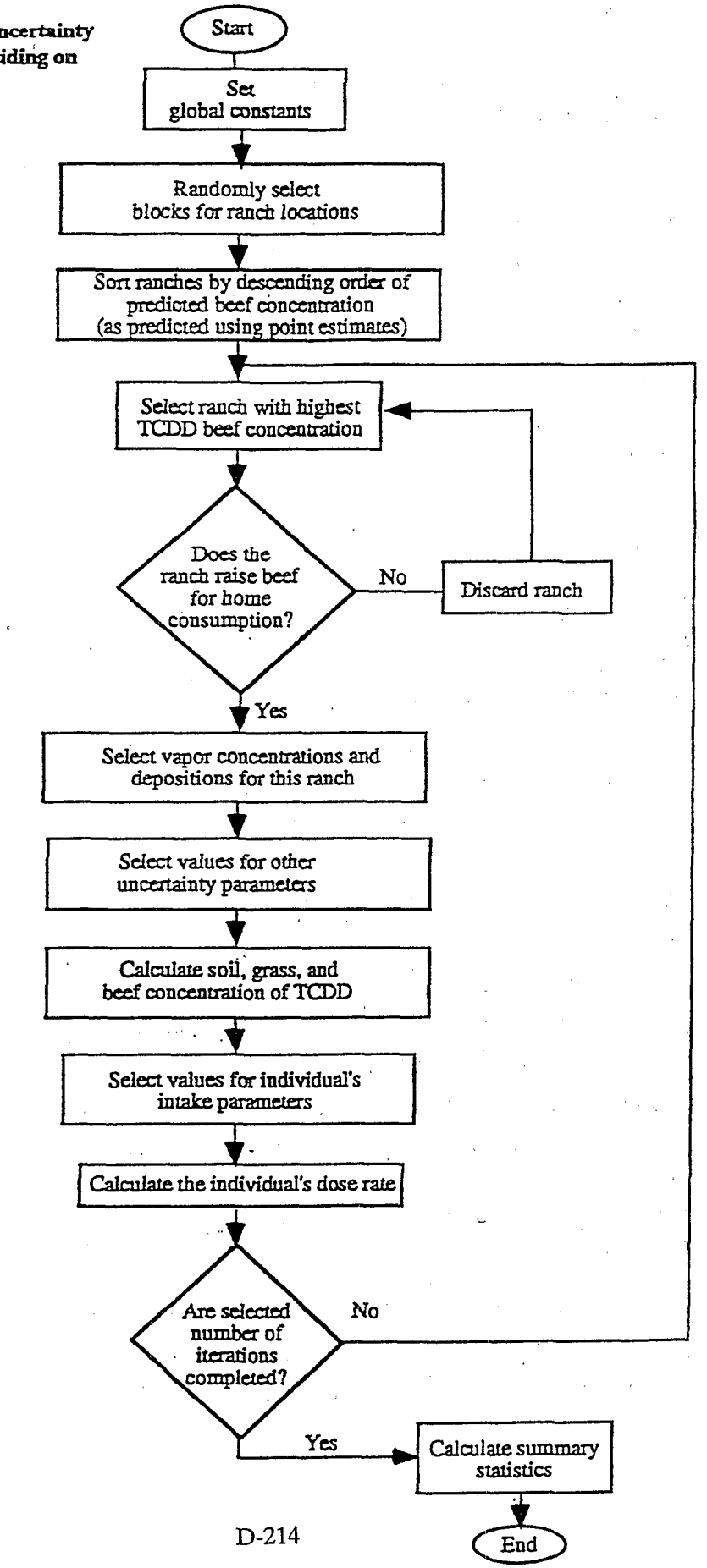
## 4.3. Monte Carlo Model of the Uncertainty in Dose Rates for the Individual Living on the Most Highly Exposed Ranch Where Beef is Raised for Home Consumption

In this analysis, the total uncertainty in the dose rate for the individual raising beef for home consumption on the ranch that receives the greatest exposure to TCDD emissions was calculated. This analysis does not differentiate between uncertainty and variation. Interpersonal variation is merely treated as another source of the total uncertainty in the individual's dose rate. Figure 2 presents a flow chart for the model.

The model selects the location of the highest exposed individual in the following manner:

1. The ranching blocks are selected in the same fashion as the local population model.
2. The blocks are then ranked in terms of their predicted beef concentration. Beef concentrations are predicted using EPA[1,6,7].
3. The model selects the ranch block with the highest predicted beef concentration and asks if cattle are slaughtered for home consumption. The probability of answering yes to this question is set at 5% (95% probability of answering no). If the answer is yes, then the airborne vapor concentration and deposition rates for this block are used in the model. If the answer is no, then the model selects the next highest block and repeats the process.

Figure 2. Monte Carlo Model of Uncertainty in the Dose Rate for Individual Residing on the Most Exposed Ranch



4. This process is repeated until a yes answer is obtained. If all of the blocks are queried without a yes answer, then the model automatically defaults to the ranch block with the lowest predicted beef concentration.

The selected values for exposure parameters are used to calculate a dose rate for the individual consuming beef raised at the selected location, and the process is repeated for 2,000 iterations. This Monte Carlo model was constructed using the same software and operates on the same platform as the previous model.

## 5. RESULTS

The point estimate analysis estimated a dose rate of $1 \times 10^{-9}$ µg/kg-day for high-end individual and a dose rate for a typical individual of $1 \times 10^{-10}$ µg/kg-day. The results of the Monte Carlo model of the uncertainty and variability in the population of individuals who live on ranches within the 400 $km^2$ area are given in Figure 3 and Table 5. Figure 3 presents the distribution of dose rates predicted for the exposed population. The separate lines in Figure 3 represent the sets of estimates for each of the dose percentiles that reflect different levels of uncertainty.

The dose rates for the different percentiles of the population range from $4 \times 10^{-13}$ to $2 \times 10^{-10}$ for the 5th and the 95th percentiles of the population, respectively (as measured by the 50th percentiles of the uncertainty distributions), or 2.5 orders of magnitude. The 90 percent confidence limits (the 5th and 95th percentiles in the uncertainty distribution of the dose rate) for the dose rate of the median individual in the exposed population range from $6 \times 10^{-12}$ to $2 \times 10^{-11}$ or slightly more than one-half of an order of magnitude. The total range of dose rates from the 90th percent lower confidence limit (LCL) of the 5th percentile of the dose rate distribution to the 90th percent upper confidence limit (UCL) of the uncertainty for the 95th percentile of the dose rate distribution range from $2 \times 10^{-13}$ to $3 \times 10^{-10}$ or approximately 3 orders of magnitude.

The total uncertainty in the dose rate for an individual randomly selected from the exposed population is given in Figure 4 and Table 6. The distribution of uncertainty for the 5th and the 95th percentiles ranges from $4 \times 10^{-13}$ to $2 \times 10^{-10}$ or about 2.5 orders of magnitude.

Figure 3. Variation and Uncertainty in Dose Rates (ug/kg-day) of the Exposed Population

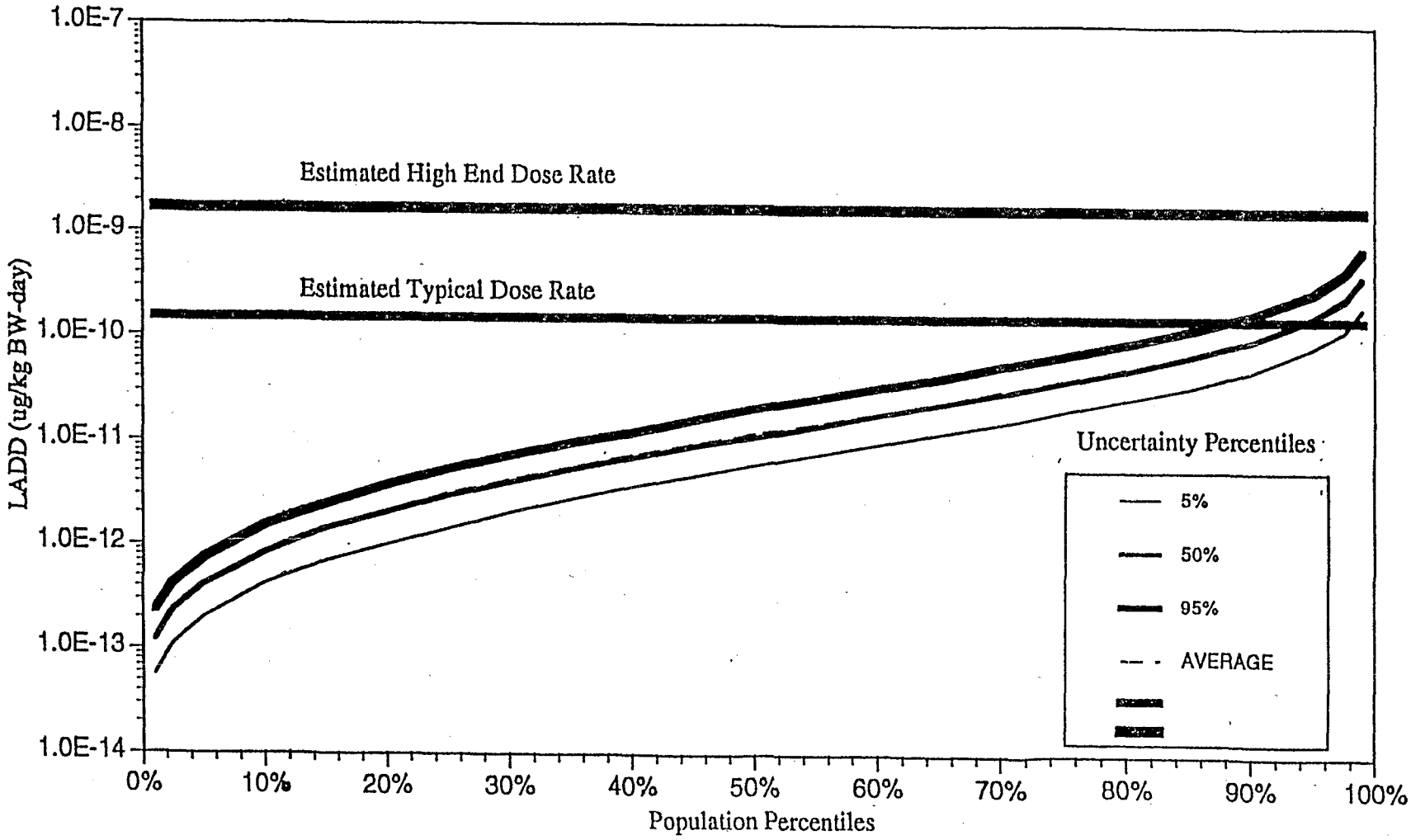


**Table 5. Uncertainty and Variation in Dose Rate (ug/kg-day) from Indirect Exposure to TCDD**

| Variance / Uncertainty | 5.0%'ile | 10.0%'ile | 25.0%'ile | 50.0%'ile | 75.0%'ile | 90.0%'ile | 95.0%'ile |
|---|---|---|---|---|---|---|---|
| 5.0%'ile | 2E-13 | 4E-13 | 2E-12 | 6E-12 | 2E-11 | 5E-11 | 8E-11 |
| 10.0%'ile | 2E-13 | 5E-13 | 2E-12 | 7E-12 | 2E-11 | 6E-11 | 1E-10 |
| 25.0%'ile | 3E-13 | 6E-13 | 2E-12 | 9E-12 | 3E-11 | 8E-11 | 1E-10 |
| 50.0%'ile | 4E-13 | 8E-12 | 3E-12 | 1E-11 | 4E-11 | 1E-10 | 2E-10 |
| 75.0%'ile | 5E-13 | 1E-12 | 4E-12 | 2E-11 | 5E-11 | 1E-10 | 2E-10 |
| 95.0%'ile | 7E-13 | 2E-12 | 5E-12 | 2E-11 | 7E-11 | 2E-10 | 3E-10 |

Figure 4. Uncertainty of Dose Rate, (ug/kg-day) Received by Individuals at a "Typical" and "Most Exposed" Ranch

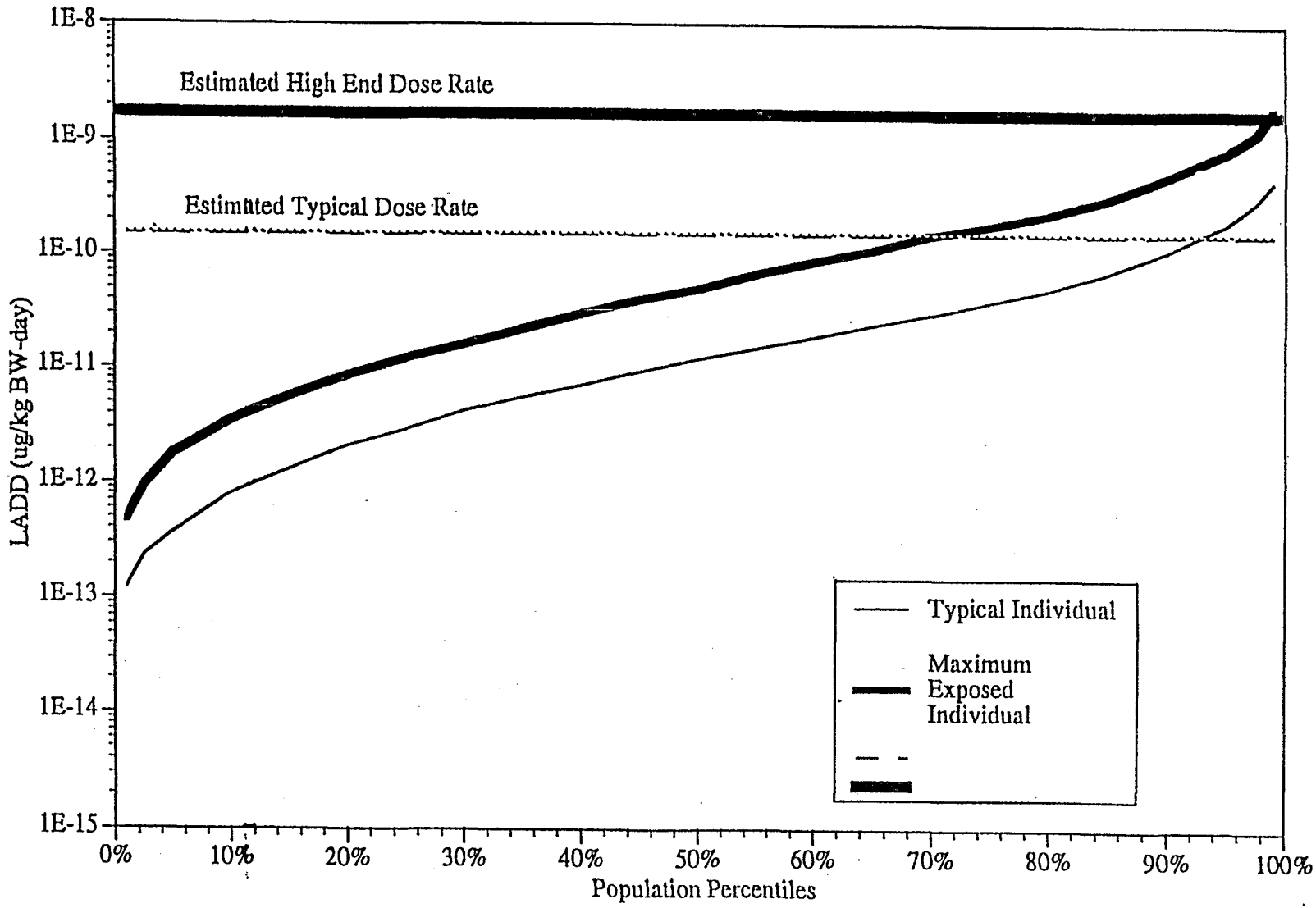


**Table 6. Uncertainty in the Dose Rate (mg/kg-day) for an Individual in the Local Population and an Individual Living on the Most Exposed Ranch (a)**

| Uncertainty | Individual Randomly Drawn from the Local Population (a) | Individual from the Ranch with Highest Exposure (a) |
|---|---|---|
| 5.0%'ile | 4E-13 | 2E-12 |
| 25.0%'ile | 3E-12 | 1E-11 |
| 50.0%'ile | 1E-11 | 5E-11 |
| 75.0%'ile | 4E-11 | 2E-10 |
| 95.0%'ile | 2E-10 | 8E-10 |

(a) That raises beef for home consumption

As discussed above, a second Monte Carlo model was conducted to evaluate the uncertainty in the dose rate for an individual who lives on the ranch that receives the highest exposure to TCDD emissions and on which beef are raised for home consumption. The results of this analysis are given in Figure 4 and Table 6. The distribution of total uncertainty for the 5th and the 95th percentiles of the dose rate ranges from $2 \times 10^{-12}$ to $8 \times 10^{-10}$ or about 2.5 orders of magnitude. In general, the dose rates received by an individual living on the most highly impacted ranch has a dose rate that is 2-3 times higher than the dose rate for a member of the general population.

## 6. DISCUSSION

The results of this analysis suggest that exposures to TCDD via consumption of beef by ranchers in a 400-$km^2$ area surrounding a typical hazardous waste incinerator could have a total uncertainty of 3 orders of magnitude. This total uncertainty is dominated by interindividual variation with uncertainty in the parameter values having a smaller effect.

The analysis confirms the observations of McKone and Ryan[22] that TCDD levels in beef can vary by several orders of magnitude, suggesting that current methods of estimating doses in potentially affected individuals have a total uncertainty of more than 3 orders of magnitude. The analysis is also consistent with McKone's[32] finding that variation is more important than uncertainty for compounds where there is direct information on the biotransfer rates. Finally, the model suggests that additional information on key parameters will not reduce the uncertainty in the dose rate received by an exposed population below 2-3 orders of magnitude because of the inherent variability in individual dose rates.

This analysis suggests that current guidance for predicting exposures to TCDD through the beef consumption pathway results in dose rate estimates greater than the 90 percent confidence limits for both the high-end and typical individual in the exposed population. The estimate of the high-end exposed individual is a factor of 4 higher than the 90 percent UCL of the 95th percentile of the general population and approximately factor of 2 higher than 90 percent UCL of the individual living on the most exposed ranch. The estimate of the typical exposed individual was 1 order of magnitude higher than the 50th percentile of the uncertainty in the dose rate of the median individual in the exposed population.

As noted in the methodology section, parameters that are not treated as point estimates are classified as representing either uncertainty or interindividual variation. Because uncertainty and interindividual variation are treated independently, the uncertainty distribution appears as a series of equally spaced bands around the dose distribution that represents interindividual variation. Future work could quantitatively consider the uncertainty inherent in characterizing the interpersonal variation distributions through either measurement error or random field modeling techniques[33,34].

Under the approach used in this paper, we assigned a specific location to the hypothetical incinerator. As a result, all the findings of this analysis are specific to the location of the hypothetical incinerator and may not be applicable to the uncertainties and variation in exposures that could occur at other locations.

There are several sources of uncertainty that are not considered in this analysis. These sources include: the uncertainty in the application of national beef consumption rates to the exposed population; the uncertainty in the diet-to-beef biotransfer factors; and the uncertainties in estimating TCDD vapor concentrations. Because the current analysis does not consider these additional sources of uncertainty, we anticipate that the actual range of TCDD dose rates in exposed individuals is greater than indicated by this analysis. In addition, because many uncertainties that have not been considered are likely to reduce rather than to increase exposures to TCDD, we believe that the dose rates estimated in this analysis represent the upper portion of the total uncertainty in dose rate estimates.

## 7. CONCLUSIONS

Nested Monte Carlo models of uncertainty and variation can provide considerable insight into the uncertainty in the range of doses potentially received by populations exposed via indirect exposure pathways. For example, this analysis suggests that while both uncertainty and interindividual variation are significant in indirect exposures, the total uncertainty is dominated by variation. The use of Monte Carlo models also allows the consideration of factors such as the likelihood of raising beef for home consumption. The analysis also suggests that the use of point estimates in indirect assessments can lead to overestimates of dose rates. Because the location of the facility was selected to maximize the potential for a beef consumption pathway, the degree of overestimation for other locations may be higher than this analysis indicates. These findings suggest that any

determination of unacceptable risk based on the use of current EPA guidance documents should be confirmed by a probabilistic analysis of the range of doses that could occur in an exposed population, and the certainty with which existing data allow that distribution to be characterized.

## ACKNOWLEDGMENT

This work was supported by the Chemical Manufacturers Association.

## 8. REFERENCES

1. Environmental Protection Agency (EPA), "Methodology for Assessing Health Risks Associated with Indirect Exposure to Combustor Emissions," U.S. Environmental Protection Agency, Office of Health and Environmental Assessment, Washington, DC. EPA/600/6-90/003. January (1990).

2. Lorbor, M., D. Cleverly, J. Scharm, L. Phillips, G. Schwear, and T. Leighton, "Development and Validation of an Air-to-beef Food Chain Model for Dioxin-like Compounds," *Sci. Tot. Environment* **156**:39-65. (1994).

3. Fries, G.F. and D.J. Paustenbach, "Evaluation of Potential Transmission of 2,3,7,8-Tetrachlorodibenzo-p-dioxin Contaminated Incinerator Emissions to Humans Via Foods," *J. Toxicol. Environ. Health* **29**: 1-43. (1990).

4. Keenan, R.E., M.M. Sauer, F.H. Lawrence, E.R. Rand, and D.W. Crawford, "Examination of Potential Risks from Exposure to Dioxin in Sludge Used to Reclaim Abandoned Strip Mines," *in* D.J. Paustenbach (ed.), *The Risk Assessment of Environmental and Human Health Hazards: A Textbook of Case Studies* (John Wiley & Sons, New York), pp. 935-998. (1989).

5. Environmental Protection Agency (EPA), "Memo to The Directors from E.P. Law, Assistant Administrator. Re: EPA's Draft Waste Minimization and Combustion Strategy and Its Implications for Superfund," U.S. Environmental Protection Agency, Office of Solid Waste and Emergency Response, Washington, D.C. May 9. (1994).

6. Environmental Protection Agency (EPA), "Addendum to the Methodology for Assessing Health Risks Associated with Indirect Exposure to Combustor Emissions," U.S. Environmental Protection Agency, Office of Health and Environmental Assessment, Exposure Assessment Group, Washington, D.C. EPA/600/AP-93/003. November. (1993).

7. Environmental Protection Agency (EPA), "Guidance for Performing Screening Level Risk Analyses at Combustion Facilities Burning Hazardous Wastes," U.S. Environmental Protection Agency, Office of Emergency and Remedial Response, Washington, D.C. April. (1994).

8. Environmental Protection Agency (EPA), "Estimating Exposure to Dioxin-Like Compounds; Review Draft," U.S. Environmental Protection Agency, Office of Research and Development, Washington, D.C. EPA/600/6-88/005B. August (1994).

9. McKone, T.E. and K.T. Bogen, "Predicting the Uncertainties in Risk Assessment," *Environ. Sci. Technol.* 25(10):1674-1681. (1991).

10. Environmental Protection Agency (EPA), "Final Guidelines for Exposure Assessment," U.S. Environmental Protection Agency, Washington, D.C. Vol 57 Federal Register No. 104. May 29 (1992).

11. Cullen, A.C., "Measures of Compounding Conservatism in Probabilistic Risk Assessment," *Risk Anal.* 14(4):389-393. (1994).

12. Environmental Protection Agency (EPA), "Superfund Exposure Assessment Manual," U.S. Environmental Protection Agency, Office of Remedial Response, Washington, D.C. EPA/540/1-88/001. April (1988).

13. Environmental Protection Agency (EPA), "Review of Draft Addendum to the Methodology for Assessing Health Risks Associated with Indirect Exposure to Combustor Emissions," U.S. Environmental Protection Agency, Science Advisory Board Indoor Air Quality/Total Human Exposure Committee, Washington, D.C. EPA-SAB-IAQC-94-009b. July. (1994).

14. Morgan, M.G. and M. Henrion, *Uncertainty: A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis.* Cambridge University Press, New York, NY. (1990).

15. Bogen, K.T. and R.C. Spear, "Integrating Uncertainty and Interindividual Variables in Environmental Risk Assessment," *Risk Analysis* 7(4): 427-436. (1987).

16. Hoffman, F.O. and J.S. Hammonds, "Propagation of Uncertainty in Risk Assessments: The Need to Distinguish Between Uncertainty Due to Lack of Knowledge and Uncertainty Due to Variability," *Risk Anal.* 14(5):707-712. (1994).

17. Frey, H.C, "Separating Variability and Uncertainty in Exposure Assessment: Motivations and Method," Session 116A: Advances in Exposure Assessment Methodology 86th Annual Meeting of the Air and Waste Management Association, Denver, Colorado. June 13-8. (1993).

18. Environmental Protection Agency (EPA), "Report on the Technical Workshop on WTI Incinerator Risk Issues," U.S. Environmental Protection Agency, Risk Assessment Forum, Washington, D.C. EPA/630/R-94//001. December. (1993).

19. Travis, C.C. and A.D.Arms, "Bioconcentration of Organics in Beef, Milk, and Vegetation," *Environ. Sci. Technol.* 22(3): 271-274. (1988).

20. Belcher, G.D. and C.C. Travis, "Modeling Support for the RURA and Municipal Waste Combustion Projects: Final Report on Sensitivity and Uncertainty Analysis for the Terrestrial Food Chain Model," Prepared under IAG-1824-A020-A1 by Oak Ridge National Laboratory, Oak Ridge, Tennessee for U.S. Environmental Protection Agency, Office of Health and Environmental Assessment, Environmental Criteria and Assessment, Cincinnati, OH. (1989).

21. Stevens, J.B. and E.N. Gerbec, "Dioxin in the Agricultural Food Chain," *Risk Analysis* 8(3): 329-335. (1988).

22. McKone, T.E. and P.B. Ryan, "Human Exposures to Chemicals Through Food Chains: An Uncertainty Analysis," *Environ. Sci. Technol.* 23: 1154-1163. (1989).

23. Jensen, D.J., R.A. Hummel, N.H. Mahle, C.W. Kocher, and H.S. Higgins, "Residue Study on Beef Cattle Consuming 2,3,7,8-Tetrachlorodibenzo-p-dioxin," *J. Agric. Fd. Chem.* 29: 265-268. (1981).

24. EarthTech, Technical Memorandum to ChemRisk. July 7. (1994).

25. Israeli, M. and C.B. Nelson, "Distribution and Expected Time of Residence for U.S. Households," *Risk Analysis* 12(1): 65-72. (1992).

26. Bidleman, T.E., "Atmospheric Processes," *Environ. Sci Technol.* 22(4): 361-373. (1988).

27. Baes, C.F., R.D. Sharp, A.L. Sjoereen, and R.W. Shor, "A Review and Analysis of Parameters for Assessing Transport of Environmentally Released Radionuclides through Agriculture," Report No. ORNL-5786. Prepared by the Oak Ridge National Laboratory, Oak Ridge, Tennessee for the U.S. Department of Energy, Washington, D.C. (1984).

28. McCrady, J.K. and S.P. Maggard, "Uptake and Photodegradation of 2,3,7,8-Tetrachlorodibenzo-p-dioxin Sorbed to Grass Foliage," *Environ. Sci. Technol.* 27: 343-350. (1993).

29. United States Department of Agriculture (USDA) "Agricultural Statistics 1992," United States Government Printing Office, Washington, D.C. (1992).

30. Environmental Protection Agency (EPA), "Report to the United States Congress on Radon in Drinking Water," U.S. Environmental Protection Agency, Office of Water, Washington, D.C. EPA-811-R-94-001. March. (1994).

31. Palisade Corp. @ Risk for PC Excel Version 1.12. Palisade Corp., Newfield, N.Y. (1994).

32. McKone, T.E, "Uncertainty and Variability in Human Exposures to Soil Contaminants Through Home-Grown Food: A Monte Carlo Assessment," *Risk Anal.* 14(4):449-463. (1994).

33. Fuller, W.A., "Measurement Error Models," John Wiley & Sons: New York. (1989).

34. Christakos, G., "Random Field Models in Earth Sciences," Academic Press, Inc.: San Diego. (1992).

35. Bacci, E., D. Calamari, C. Gaggi, and M. Vighi, "Bioconcentration of Organic Chemical Vapors in Plant Leaves: Experimental Measurements and Prediction," University of Siena: Siena, Italy. (1989).

36. Lyman, W.J., Chapter 1: Octanol/water Partition Coefficient. In: Handbook of Chemical Property Estimation Methods. W.J. Lyman, W.F. Reehl and D.H. Rosenblatt (eds.), New York: McGraw-Hill Book Company. 1-1 to 1-47. (1982).

37. USDA, "U.S. Department of Agriculture. Household Food Consumption Survey, 1965-1966," Report 12. Food Consumption of Households in the United States - Seasons and Year, (1965-1966).

38. Marple, L., B. Berridge, and L. Throop, "Measurement of the Water-octanol Partition Coefficient of 2,3,7,8-Tetrachlorodibenzo-p-dioxin," Environ. Sci. Technol. **20**(4): 397-399 (1986).

39. Mackay, D., W.Y. Shiu, and K.C. Ma, Illustrated Handbook of Physical-Chemical Properties and Environmental Fate for Organic Chemicals; Volume II: Polynuclear Aromatic Hydrocarbons, Polychlorinated Dioxins, and Dibenzofurans. Lewis Publishers, Chelsea, MI. (1991).

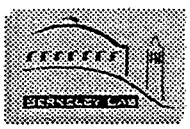
40. Schroy, J.M., F.D. Hileman, and S.C. Cheng, "Physical/Chemical Properties of 2,3,7,8-Tetrachlorodibenzo-p-dioxin," In: Aquatic Toxicology and Hazard Assessment: Eighth Symposium, ASTM STP 891. R.C. Bahner and D.J. Hansen (eds.), Philadelphia, PA: American Society for Testing and Materials. 409-421. (1985).

# PRESENTING RESULTS

**Thomas E. McKone**
**University of California**
**Berkeley, California**

# Monte Carlo Issues Workshop

## Presenting Results



Thomas E. McKone, Ph.D.
University of California
Ernest Orlando Lawrence Berkeley National Laboratory
and
University of California, Berkeley
School of Public Health

## Presenting Results

- (1) What should be presented?
- (2) How can variability/uncertainty be characterized?
- (3) How to compare Monte Carlo results to point estimates
- (4) How to characterize the results of a sensitivity analysis

## Presenting Results (continued)

- (5) How to characterized the stability of the tails
- (6) How to present the results of a expert elicitation
- (7) Incompatibility between exposure estimates and dose-response metrics
- These are questions and issues--full answers are not yet available

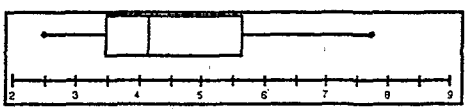
## (1) What Should be Presented?

- What elements or information must be present in a report that presents results of a Monte Carlo Analysis?
- How can the results be checked to assess quality control?
- Examples

# What to Include in a Report

- Measures of central tendency
  - » Mean, median, mode, etc.
  - » range (which?)
- Measures of spread
- Other measures
  - » Skewness, Kurtosis
  - » How to represent mixed distributions?
  - » Maximum value
  - » Confidence intervals

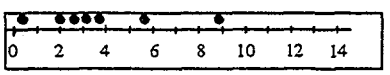
# Examples




Box and whiskers plot

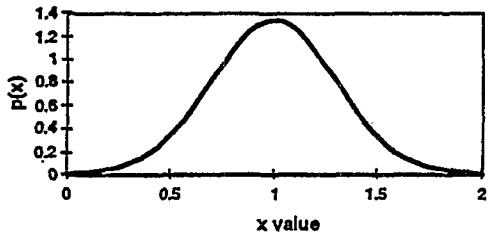
Stem and leaf plot

| | |
|---|---|
| 1 | 0,0,1,3,4,5,6,6,7,8 |
| 2 | 2,4,6 |
| 3 | 0,1,5,8 |
| 4 | 0,0,3,4,5,8,9 |
| 5 | 2,3,6 |
| 6 | 0 |
| 7 | 3,5 |
| 8 | 0,2,5 |
| 9 | 1,4,5 |
| 10 | 0 |



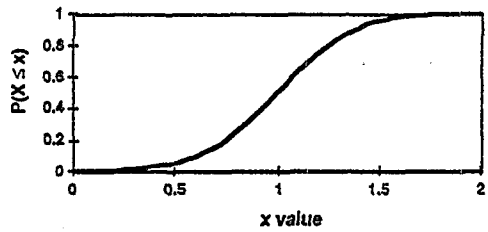

Dot plot

# Examples

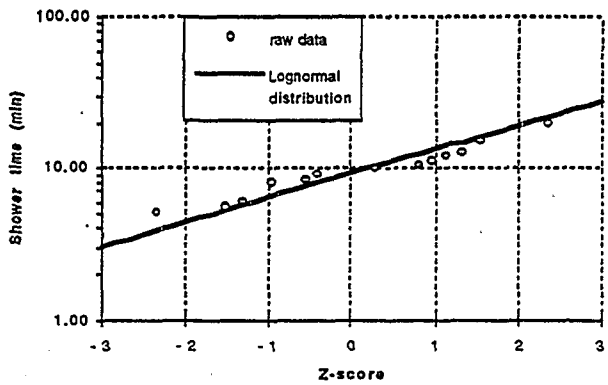



PDF

CDF




# Examples




**Probability plot using Z scores**

## (2) Characterizing Variability and Uncertainty

- Define and illustrate methods for characterizing variability and uncertainty in model inputs
- Identify ways in which models are modified to accommodate both variability and uncertainty

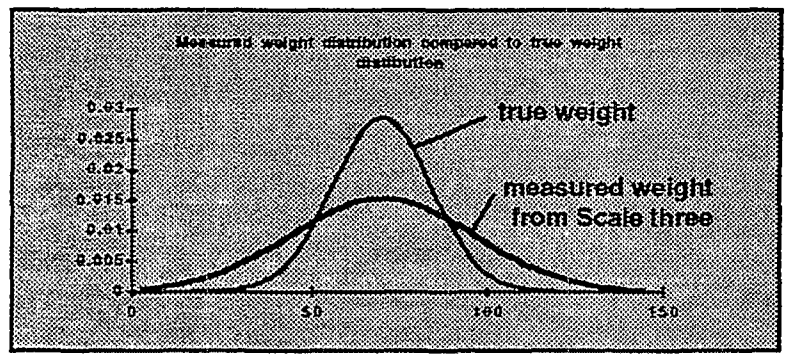
## Variability/Uncertainty Example

- Measure the distribution of body weight, BW, in a population with:

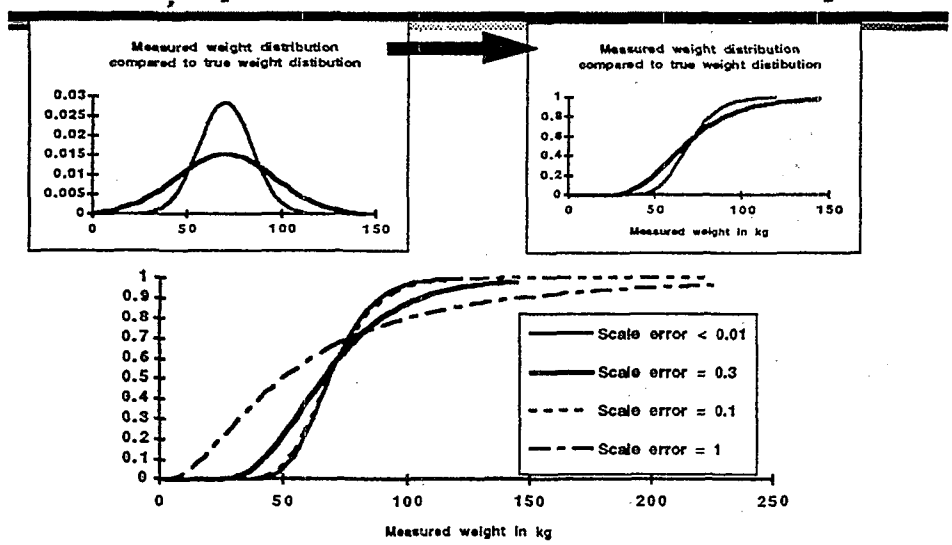
Mean BW = 70 kg  Stdev = 14 kg

# Four Scales of Differing Precision (CV)

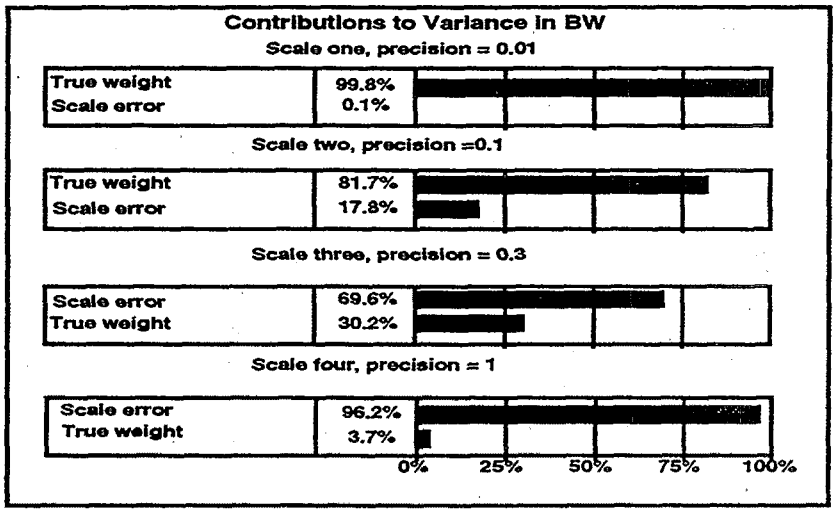
- Scale one = 1%  Scale two = 10%
- Scale three = 30%  Scale four = 100%




# Scale precision and uncertainty



## Sensitivity Chart for BW Measurements



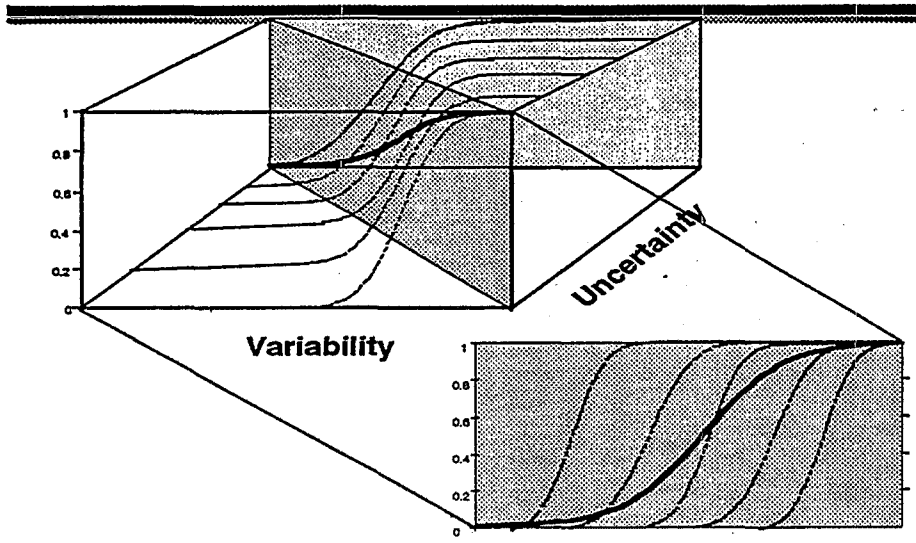

## Separating Uncertainty and Variability

$$\text{Measured body weight} = \text{Actual body weight} \times \text{Scale error} \times [\text{Scale Bias}]$$

Variability — Uncertainty

**In our example the bias is zero**

## 3-D View




## (3) Monte Carlo results versus Point Estimates

- What is the best way to compare point estimates of exposure to the output of a Monte Carlo simulation?
- What information should be included in the discussion of this comparison?
- Enumerating the benefits and limitations of Monte Carlo
- What discussion is needed when the Monte Carlo results differ significantly from the point estimate?

## (4) Characterizing the Results of a Sensitivity Analysis

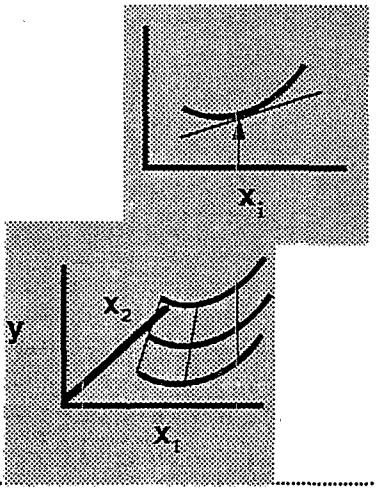
- How to characterize and discuss results
- How to characterize the influence of the sensitivity analysis on the selection of point estimates or distributions for inputs
- Examples

## Examples: Key Sources of Variability and Uncertainty

- This topic is iterative in nature.
- Identify key contributors to variability and uncertainty *before* doing a 2-D analysis (to simplify the problem)
- Determining if uncertainty or variability dominates, so that a 2-D analysis is not required

## Sensitivity Analyses

- Rate of change of output with respect to change in input
- Local sensitivity about the nominal value
- Global sensitivity over the entire parameter space




## Identify the Most Sensitive Inputs

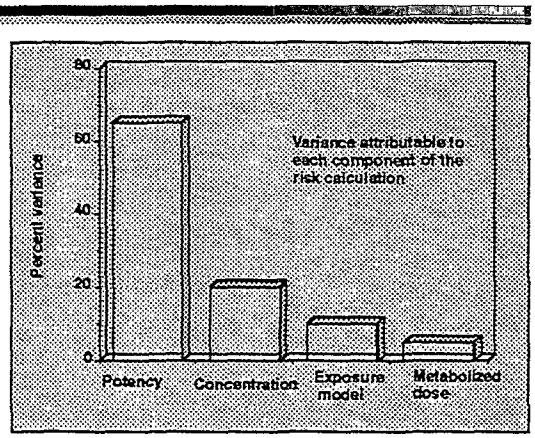
- Additional data collection
- Additional research
- Stratification of the population

# Analysis of Variance

- An uncertainty analysis provides a way to determine the amount of outcome variance attributable to specific input parameters or groups of input parameters.
- There are several methods for ranking uncertainty, including
  - » correlation coefficients and regression coefficients,
  - » rank correlation and rank regression.
- These methods are designed to rank the contribution of individual parameters.

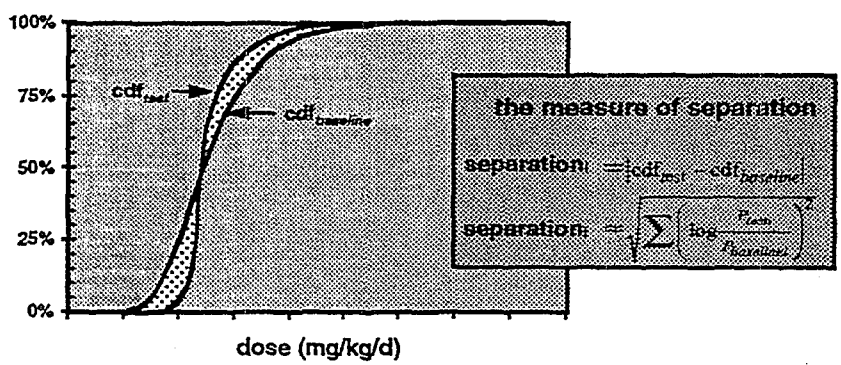
# Attributable Variance

- One approach that makes possible the ranking of either individual parameters or groups of parameters involves the use of repeated Monte Carlo simulations with one or more parameters excluded from the analysis.



# Selection Criteria

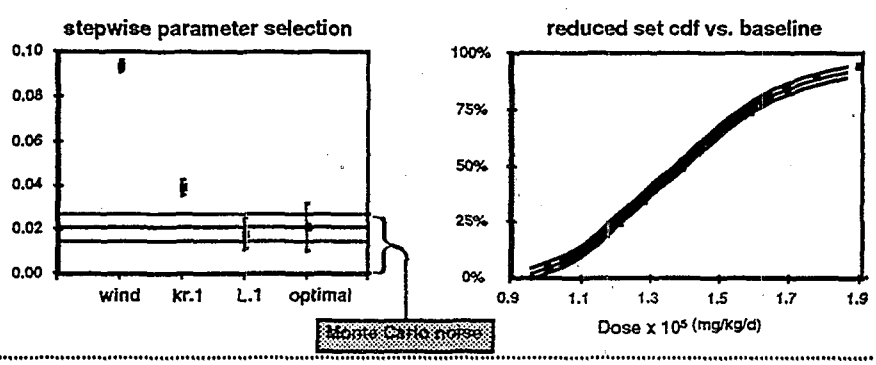
- **Curve separation**
  - » absolute distance between test cdf and "baseline"




# Selection Process and Results

- **Case study: Benzene**

  Pass 1: select wind speed
  Pass 2: include reaction rate (air)
  Pass 3: include depth (air) - separation < noise (STOP)
  Pass 4: optimize by removing non-influential parameters



## (5) Characterizing the Stability of the Tails

- How can one adequately characterize the stability of the tails of the output distribution?
- Can adequate discussion of the confidence in the "high-end" values be provided?

## (6) Presenting the Results of a Expert Elicitation

- When profession judgment or Delphi techniques are used, what is the best way to describe the process in the presentation of results?
- Which factors of the decision should be listed?
- How can the importance of the variations be characterized?

## (7) Matching Exposure Estimates to Dose-Response Metrics

- Because of fixed exposure assumptions imbedded in toxicity metrics, the output distribution of exposure estimates may be incompatible with the dose-response endpoint selected for risk assessment.
- How can this situation be avoided
- Should there be probability distributions for the dose-response values?

# COMMUNICATING AND DOCUMENTING UNCERTAINTY IN RISK ANALYSIS

**Max Henrion**
**Lumina Decision Systems, Inc.**
**Los Altos, California**

## Some sources of uncertainty for the reviewer of a risk analysis

- For parameter uncertainties, what distributions are used and why?
- What is the model uncertainty?
- What issues and factors does the model include?
- What assumptions does the model make?
- Are the model equations implemented correctly?

Lumina DECISION SYSTEMS

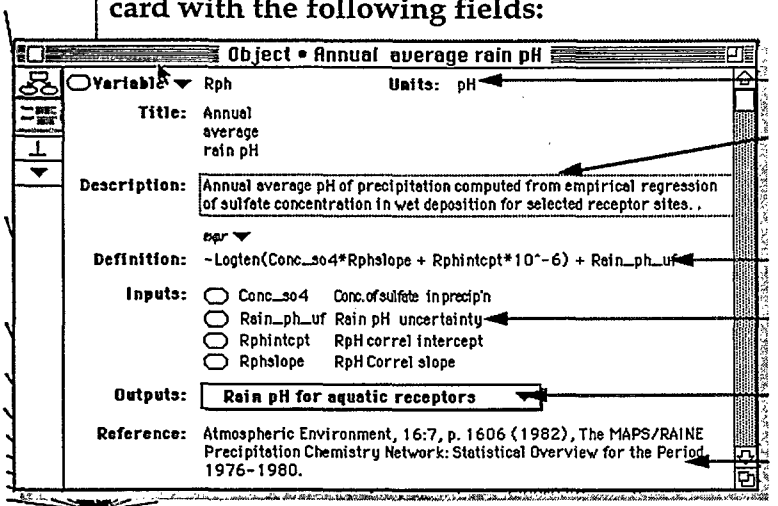
## The importance of communicating and documenting risk analyses

- Uncertainty analysis is about assessing and communicating the appropriate degree of confidence in the results
- Clarity in communicating model assumptions and structure is as important as clarity in communicating uncertainty about results
- The presentation of the risk analysis should be designed to facilitate QA and review
- Large spreadsheets and proprietary computer codes can be obstacles to clarity

Lumina DECISION SYSTEMS

# Integrated model documentation

**Each variable and module is documented by a card with the following fields:**




Units of measurement

What it represents

Math. formula for calculation

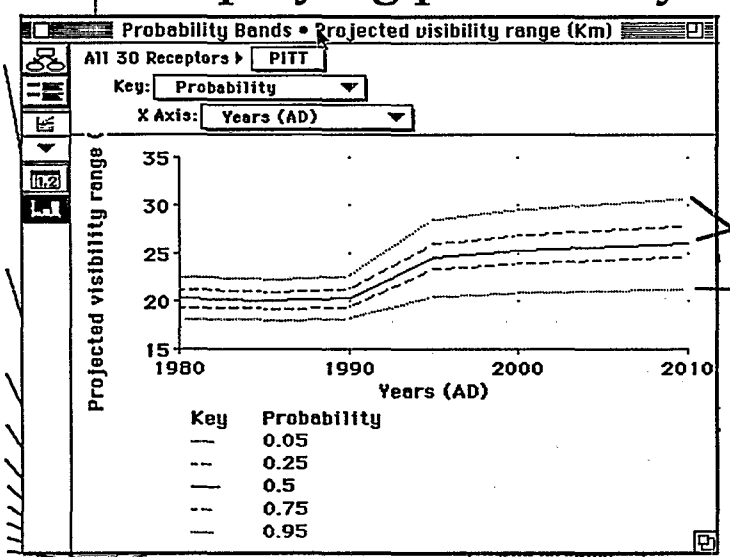
What it depends on

What depends on it

Source or citation

Lumina Decision Systems

# Displaying probability bands

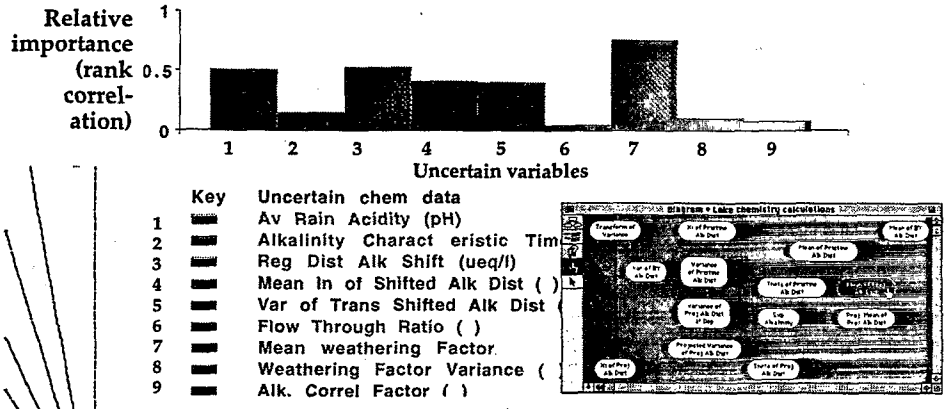



median

90% probability interval

Lumina Decision Systems

# Uncertainty analysis to identify key sources of uncertainty




Analysis of the relative importance of the uncertainty in each of the input parameters of the lake-chemistry model to the uncertainty in the fraction of lakes with pH < 5.5

Lumina DECISION SYSTEMS

# Summary

- **Clarity in communicating model assumptions and structure is as important as clarity in communicating uncertainty about results**
- **Three helpful tools are:**
  - **Hierarchical influence diagrams**
  - **Integrated model documentation**
  - **Sensitivity analysis to identify the relative importance of sources of uncertainty**
- **The software model should be delivered and reviewed along with the paper report**

Lumina DECISION SYSTEMS

# EIGHT REASONS TO CONSIDER UNCERTAINTY

**Max Henrion**
**Lumina Decision Systems, Inc.**
**Los Altos, California**

# Eight reasons to consider uncertainty


## Max Henrion

**Lumina Decision Systems, Inc.**

4984 El Camino Real, Suite 105

Los Altos, CA 94022

Voice 415-254-0189 Fax 415-254-0292

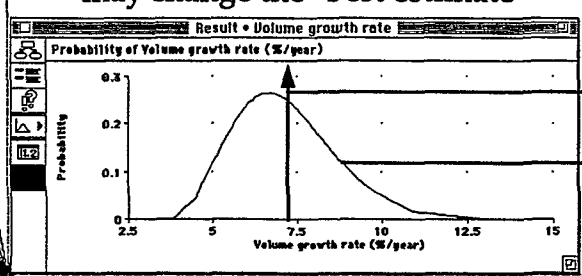
Internet: henrion@lumina.com

WWW: http://www.Lumina.com/Lumina


Lumina DECISION SYSTEMS

## Why model uncertainty? 1 to 4

1. Because you are risk averse
2. Because you want to decide what information to buy
3. Because your results will need to be combined with other sources of information
4. Because thinking about the uncertainty may change the "best estimate"




Certain or uncertain?

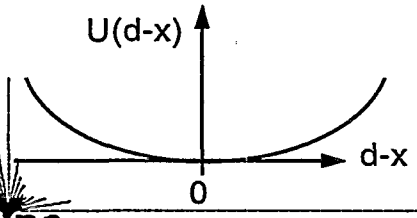
Lumina DECISION SYSTEMS

# Why model uncertainty? 5

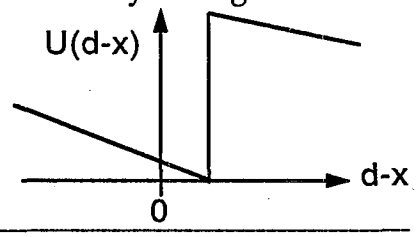
5. Because it may change the recommended decision leading to a higher expected value relative to ignoring uncertainty.

   The Expected Value of Including Uncertainty (EVIU) is the increase in expected value when you model uncertainty explicitly instead of assuming all uncertain quantities are certain to be at their mean value.

If utility is linear or quadratic in x, EVIU=0




ility has a step function, EVIU may be large




Lumina DECISION SYSTEMS

# Why model uncertainty? 6 to 7

**6. Because it's more ethical to acknowledge the limitations of your analysis**

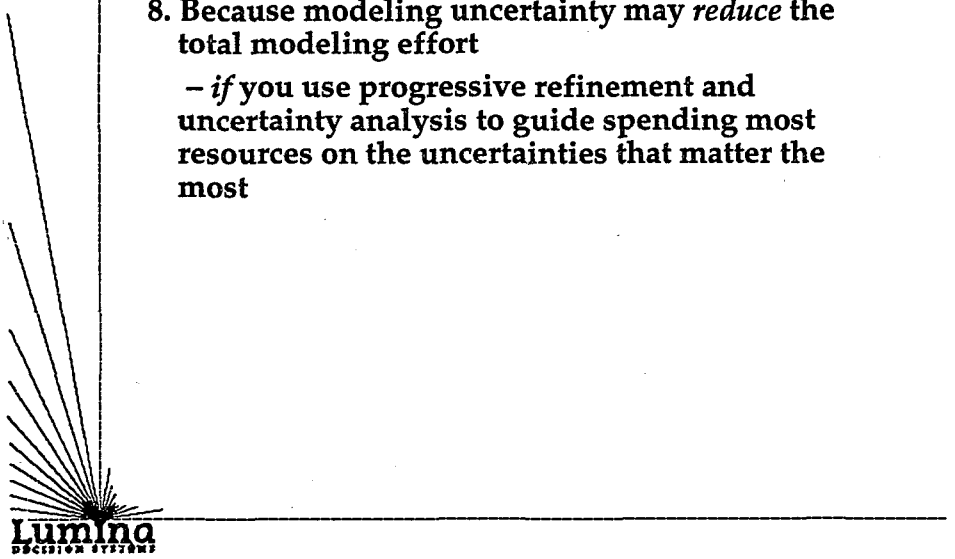
**7. Because it's safer to admit you're uncertain**

*"God created the World in 4028 BC on the ninth of September at nine o'clock in the morning"*
Archbishop Ussher of Ireland AD 1658.

Lumina DECISION SYSTEMS

# Why model uncertainty? 8

8. Because modeling uncertainty may *reduce* the total modeling effort
   - *if* you use progressive refinement and uncertainty analysis to guide spending most resources on the uncertainties that matter the most



# APPENDIX E

# STATEMENTS DEVELOPED BY THE WORKGROUPS

## A TIERED APPROACH TO UNCERTAINTY/VARIABILITY ANALYSIS IN EXPOSURE ASSESSMENT

Workgroup Chair: Tom McKone

Health-risk assessments provide quantitative evaluations of the potential health hazards of various agents and on the extent of human exposure to these contaminants. Risk assessments involve four interrelated steps: 1) hazard identification, 2) dose-response characterization, 3) exposure assessment, and 4) risk characterization. Many sources of both uncertainty and variability are present in each of these steps. Effective environmental management policies are possible under conditions of both uncertainty and variability, but such policies must take both into account. In this section, we considered how uncertainty and variability can be confronted in the exposure component of health risk assessment. We describe a tiered approach to confronting uncertainty that is flexible and allows for a smooth transition from the existing risk-assessment framework that addresses uncertainty by quantifying a plausible upper bound on exposure and risk to a framework in which the risk and exposure are characterized with distributions that reflect uncertainty and variability.

### The Nature and Quality of Information Used in Exposure Assessment

Exposure assessments contribute to a number of health-related assessments, including risk assessments, status and trends analyses, and epidemiological studies. In many exposure scenarios, such as those for combustion sources or contaminated soils and ground water, a multimedia approach is needed. In contrast to the single-medium paradigm, which typically has been used for assessing exposure, in a multimedia approach we locate all points of release to the environment, characterize mass-balance relationships, trace contaminants through the entire environmental system, define changes in chemical form as they occur, and identify where in this chain of events control efforts are likely to be most effective. Multimedia-exposure models, however, require that we measure or estimate transfer coefficients of contaminants among two or more environmental media and among environmental media and exposure media. This can increase the level of uncertainty in exposure predictions.

Defining exposure pathways is an important component of the exposure assessment. An exposure pathway is the course a chemical, biological, or physical agent takes from a known source to an (often unknown) exposed individual. An exposure pathway describes a unique mechanism by which an individual or population is exposed to chemical, biological, or physical agents at or originating from a source. Each exposure pathway includes a source or release from a source, an exposure point, and an exposure route.

The data, scenarios, and models used to represent human exposures to environmental emissions include at least five important relationships that involve uncertainty and variability:

- The magnitude of the source medium concentration, that is, the level of contaminant that is released to air, soil, or water or the level of contamination measured in or estimated in the air, soil, plants, and water in the vicinity of the source.
- The contaminant concentration ratio, which defines how much a source-medium concentration changes as a result of transfers, degradation, partitioning, bioconcentration, and/or dilution to other environmental media before human contact.
- The level of human contact, which describes (often on a body-weight basis) the frequency (days per year) and magnitude (kg/day) of human contact with a potentially contaminated exposure medium.
- The duration of potential contact for the population of interest as it relates to the fraction of lifetime during which an individual is potentially exposed.
- The averaging time for the type of health effects under consideration, that is, the appropriate averaging time for the cumulative duration of exposure such as a human lifetime (as is typical for cancer and chronic diseases) or some relatively short time period (as is the case for acute effects).

These factors typically converge as a sum of products or quotients to define a distribution of population exposure or a range of individual exposures. Typically, model specification errors are not significant in the exposure models. Thus, any expected variance in estimates of exposure is attributable to mainly to input variance. The relationship between the variance of population exposure estimates and input variances, which reflect uncertainty and variability in these above factors, can be determined using a tiered approach to uncertainty/variability as described below.

**A Tiered Approach**

An important, and often ignored, final step in the exposure and risk characterization process is the characterization of uncertainties. This process is implicit to the premise behind the risk-based approach, but frequently passed over in actual practice. In order to more directly confront uncertainties in risk assessments, it is necessary to take a tiered approach to uncertainty analysis. We identify here a four-tiered approach that includes progressively increasing levels of complexity. Our intent is to suggest a range of sophistication and not to define a comprehensive and rigorous process. In applying this tiered approach, the level of effort should be commensurate with the scope of the problem. For example, an comprehensive stochastic model for soil contamination may not be needed to assess the health risks of contaminants that are deposited on soil from the atmosphere.

There are five factors that determine the precision or reliability of a health impact assessment: 1) specification of the problem (scenario development), 2) formulation of the conceptual model (the influence diagram), 3) formulation of the computational model, 4) estimation of parameter values, and 5) calculation and documentation of results including uncertainties. An uncertainty analysis involves the determination of the variation or imprecision in an output and how this relates to variance of model inputs and/or to model error or problem specification error.

The tiered process as discussed below should apply separately to all relevant subpopulations and to all relevant scenarios (i.e., occupational versus residential, short-term versus long-term, alternative land use, accidental versus routine, etc.). We view these steps as sequential. One should not begin at tier four without having worked through the first three tiers. However, the tiers described below are not intended to be prescriptive and we emphasize that the level of effort should be commensurate with the scope of the problem.

**Tier 1: Single-Value Estimates of High-End and Mid-Range Risk**

The first tier in a flexible approach to uncertainty involves straightforward point estimates of risk to a high end individual (e.g., bounding estimate of exposures) and to the mid-range individual

(representing the more typical member of a given population or subpopulation) with discussion of how plausible and likely is the high end and the mid-range. If the point estimate of high-end risk is lower than the regulatory level of concern, the risk analysis may stop here. It is useful at this level to include at a minimum some qualitative or quantitative evaluation of what the most uncertain and most sensitive parameters are and how this affects the result. This permits interpretation and evaluation of the results.

**Tier 2: Qualitative Evaluation of Model and Scenario Sensitivity**

At the second tier of complexity, exposure scenarios, conceptual model formulation, and the computational model structure, including the selection of parameter combinations, should be evaluated to determine how alternate models and/or scenarios impact the bounding (high-end or mid-range) estimates of exposure or risk. The purpose of this step is to help identify errors or omissions in the development of the conceptual model. At a minimum, this can be done by listing the estimation error, the experimental variance, and the population variability ranges associated with parameters in the written reports where these parameters or their estimation equations are defined. A clear summary and justification of the assumptions used for each aspect of a model, stating whether these assumptions are likely to result in representative values or conservative (upper bound) estimates, help to define and reduce uncertainties. At this tier, if alternate models for representing transport, dispersion, uptake, etc., were considered, there should be an effort to consider the impact of alternative models on the magnitude of the bounding exposure estimates. If this exercise reveals a need to revise the analysis of the upper-bound point estimate of risk, then the updated estimate should be compared to some regulatory level of concern. If the revised point estimates are below the level of regulatory concern, additional analysis may not be necessary.

**Tier 3: Quantitative Sensitivity Analysis of High-End or Mid-Range Point Estimates**

At the third level of complexity, a sensitivity analysis should be used to assess how predictions of exposure are affected by both model reliability and the quality of the input data. The goal of a

sensitivity analysis is to rank the input parameters on the basis of their contribution to variance in the output. At this tier, a sensitivity analysis should be quantitative and formal. The formal sensitivity analysis is used to assess how variations of individual model parameters affect the magnitude of model results.[8] When evaluating the reliability and sensitivity of mid-range and upper bound point estimates, it is useful to calculate how a small (i.e., 1 percent) change of an input value affects a bounding point estimate (i.e., 1 percent or greater). This measure of change is the sensitivity of the model about a single reference outcome and is an approximation of partial derivatives used to construct a linear approximation. Multiplying the point sensitivity value by the coefficient of variation of a parameter is a useful way to get an approximate normalized rank of the importance of each parameter's uncertainty to the magnitude of uncertainty associated with the point estimate The quantitative uncertainty analysis process in Tier 3 is simply a more quantitative way to elicit errors in model framework or uncover ways in which the model has been incorrectly specified. If the estimated risk remains higher than the level of regulatory concern, fourth-tier methods should be used to get a better characterization of the problem and more insight on how to develop an intervention strategy.[8]

**Tier 4: Fully Quantitative Characterization of Uncertainty and Uncertainty Importance**

At the fourth tier, variance propagation methods (including but not necessarily Monte Carlo methods) should be used to carefully map how the overall precision of risk estimates is tied to the variability and uncertainty associated with the models, inputs, and scenarios. Mapping the uncertainty of input parameters into the uncertainty of an output variable involves the following steps: 1) identify the inputs that contribute significantly to model prediction uncertainty (this should have been completed as part of the Tier 3 process), 2) construct a probability density function (PDF) for each input that defines the range of values an input parameter can fall within and reflects the likelihood that the parameter will take on the various values within that range, 3) account for correlations (dependencies) among the inputs, 4) propagate the uncertainties through the model to generate a PDF of predicted outcome values, and 5) derive confidence limits and

[8]See "A Hierarchy of Methods for Sensitivity Analysis" (pp. E-10 - E-15) for a description of commonly used sensitivity-analysis methods and additional discussion on methods for ranking the importance of individual parameter uncertainty.

intervals from the predicted outcome variable PDF in order to provide a quantitative statement about the effect of input uncertainty on the model predictions. Ranking uncertain parameters (as part of a parameter uncertainty analyses) helps determine each parameter's contribution to uncertainty in model predictions and can provide guidance for additional research efforts. Methods available for parameter uncertainty ranking often depend on the type of variance propagation used in characterizing the uncertainty in model outputs.[8]

## USE OF NUMERICAL EXPERIMENTS IN MONTE CARLO ANALYSIS

Christopher Frey and Scott Ferson

For Monte Carlo analyses, numerical experiments are often used to:

- **Evaluate the effect of alternative assumptions regarding probability distributions for model inputs.** For example, when preserving the moments of the distribution, evaluate the effect of plausible alternative distribution shapes on the estimate of the assessment endpoint (e.g., 95th percentile of the population).[9]

- **Evaluate the effect of alternative correlation structures among model inputs.** This could include the sensitivity of results to different values of linear correlation coefficients as well as the implications of correlation matrices for groups of model inputs.

- **Evaluate the implications of alternative models on the estimate of the assessment endpoint.** This can be done through alternative cases, in which each model is run and the results are compared (preferred), or probability trees, in which weights are assessed for each model and then all possible combinations of the models are evaluated as part of a probabilistic analysis. It is difficult, however, to estimate the weights that should be given to the models and it can be misleading to lump models together if they in fact lead to different conclusions and are based upon substantially different theoretical bases.

- **Confirm the results of other sensitivity analysis techniques.** For example, it is common to use rank correlation coefficients to identify the strength of monotonic relationships between input variables and the model output. Such analysis, however, may fail to account for shifts in the central tendency of the model result due to skewed model inputs that have little contribution to the variance in the model output and may also fail to account for effects at the tails versus the central tendency of the model result. As an alternative to analysis of covariance of model inputs and outputs, one should run the model "with" and "without" various input distributions or groups of input distributions. For example, if a regression analysis indicates that many of the probabilistic model inputs contribute little to overall variability or uncertainty, then these inputs could be set to their central values (e.g., mean, median, mode) and the analysis repeated. If the probabilistic results differ insignificantly from the base case, then it is confirmed that the model result is insensitive to that particular input. This process can be done for groups of variables as well.

---

[9]Hoffman, Owen. Presentation at the 1995 Society for Risk Analysis Annual Meeting.

## A HIERARCHY OF METHODS FOR SENSITIVITY ANALYSIS

Workgroup Chair: Mitchell Small

Three broad classes of sensitivity analysis methods can be identified ranging from simple to more complex:

- Methods that compute the model's direct response to changes in input values or assumptions. These methods generally involve simple perturbations of the model and are usually employed prior to more sophisticated evaluations of uncertainty.
- Methods conducted as part of the uncertainty analysis, often through postprocessing of the simulation results.
- Decision-driven methods that assess the impact of uncertainty in input assumptions on pending decisions and the potential loss (i.e., costs and benefits) associated with them.

A hierarchy of methods for sensitivity analysis is thus suggested.

### Level 1: Sensitivity Analyses

The simplest direct response methods explore changes in model output for a discrete or unit change in each of the inputs, one at a time. With *range sensitivity*, the input is varied from its nominal value to low and high plausible values and the model response noted. In *differential* (a.k.a. *point* or *local*) sensitivity, the partial derivative of the output with respect to the input is computed at the point of the nominal input. This can be done analytically using methods of calculus when the model is simply based on analytical equations; more generally, it is computed numerically with very small perturbations in the input around its nominal value. A *normalized differential sensitivity* may be computed by dividing the change in the model output in the numerator by its nominal value, and the change in the model input in the denominator by its nominal value, yielding a dimensionless quantity that can be more readily compared across inputs. In each case, all other inputs to the model are held at their nominal (a.k.a. baseline,

"best–guess," or point) values when the sensitivity is computed; these nominal values generally correspond to the means or medians in the subsequent probabilistic analysis.

While the simple methods described above provide a first indication of the sensitivity of the model output to each of the inputs, this picture can be clouded by failure to account for model behavior over the full range of the model input, which can be nonlinear and in some cases nonmonotonic; failure to account for interactions between parameters; and failure to consider how the model sensitivity to an input combines with the uncertainty in the input to determine the overall contribution to model output uncertainty. The former two concerns can be addressed with more sophisticated numerical studies of the model input-output space (also called the *response surface*). The latter is addressed by conducting the sensitivity analysis in the context of the overall uncertainty analysis with the second class of methods, described below. While the simple (pre-uncertainty) sensitivity methods thus provide only a partial picture from which to judge parameter importance, they can be used to screen or prioritize parameters for the subsequent uncertainty analysis. Parameters eliciting a very small response from the model can in certain cases be eliminated from further probabilistic treatment, assuming important higher-order, nonmonotonic or parameter interaction effects are also precluded.

**Level 2: Sensitivity Analyses**

The second class of methods for identifying parameter importance combine sensitivity of the model output to the parameter with the *uncertainty* in that input. The simplest of these, first-order uncertainty analysis (FOUA), approximates the output variance as the partitioned sum of the variances contributed by each uncertain input (Cox and Baybutt, 1981):

$$\sigma_y^2 = \sum_i \left(\frac{dy}{dx_i}\right)^2 \sigma_{x_i}^2$$

where dY/dXi is the differential sensitivity of the model output Y to input Xi, and Var[Xi] is the uncertainty associated with this input. Each term represents the contribution of input Xi to the overall variance, so the importance of each input is indicated by this term divided by the total (i.e., the fraction of variance in Y explained by uncertainty in Xi). The FOUA equation assumes uncertainties in the inputs are independent and that the uncertainties are small relative to nonlinearities and interactive parameter effects in the model response surface. Elaborations of FOUA (e.g., Morgan and Henrion, 1990) allow for covariance terms for correlated input uncertainty and higher-order derivatives which can address some nonlinearities, at least locally (e.g., *second*-order uncertainty analysis, which also uses the second derivative of the response surface). As a first approximation which requires only estimates of the variances of the inputs rather than specification of their full distribution functions, FOUA can be used to screen parameters for further numerical (or Monte Carlo) uncertainty analysis.

Once a full numerical Monte Carlo/probabilistic uncertainty analysis is conducted (henceforth referred to as MC/P analysis, but also including numerical methods based on nonrandom sampling, such as LHS), a suite of methods is available for assessing parameter importance, involving post-processing or modification of the numerical simulations. Unfortunately, while MC/P analysis can provide a more accurate derivation of model output uncertainty than can the approximate FOUA and related methods, interpretation of results in terms of partitioned variance is less direct. The most common of the methods rank importance based on correlation, partial correlation, or partial rank-order correlation of the model inputs with the output, the latter generally being most robust and popular. Higher correlations imply greater importance. These methods should be accompanied by an input-output scatter plot of the simulation to check for the presence of nonlinear or nonmonotonic behavior that may not be apparent from the value of these correlation measures.

A second set of methods for identifying parameter importance in the context of an MC/P uncertainty analysis involves rerunning the simulation by selectively setting the target input or set of inputs to their nominal value, thereby eliminating the probabilistic character of these inputs. The magnitude of the change in the simulated output distribution is then noted. More important parameters result in more dramatic changes in the output distribution (generally reductions in spread or variance, but also possible changes in central tendency). This method can also be used

to test the sensitivity of the model for conceptual or model structural assumptions, again noting the change in the output distribution resulting from the new or alternative formulation.

While direct and exact determination of the contribution of each uncertain parameter to output variance cannot be accomplished with MC/P methods, approximations are available. If the model can be well-represented by a linear response surface, partial correlation coefficients (squared) of the inputs approximate the contribution to variance. In certain cases, analysis of variance techniques combined with selective addition or subtraction of uncertain terms can be used. Special sampling methods such as Fourier sampling (e.g., McRae et al., 1982), are directly amenable to the partitioning of variance, but these may imply overly restrictive assumptions on the shape of the input distributions and require special computational tools.

**Level 3: Sensitivity Analyses**

The third and perhaps most advanced set of methods for assessing parameter importance measures the impact of parameter uncertainty on the pending decision(s). A first approach, requiring little additional effort beyond MC/P analysis, involves portioning the model output simulations into classes, depending on the risk management decision they imply. If the distributions of the sampled inputs in these partitioned classes differ from each other significantly, then the parameter is inferred to be important relative to the decision. This approach is similar to the sensitivity analysis methods developed by Hornberger and Spear (1980; Beck, 1987), who divided input distributions into classes depending on whether the output was "acceptable," i.e., consistent with observations of the output, except here the emphasis is the impact on decision, rather than consistency with observation. Both methods can of course be used productively at different phases of the assessment. Mertz et al. (1992) illustrate uses of logistic regression to demonstrate which parameters most dramatically affect a decision for the case of a yes-no (or zero-one) decision (e.g., whether or not to regulate, remediate).

The most general and broadly applicable set of decision–driven techniques for evaluating parameter importance involve the calculation of value-of-information (VOI), or data-worth (e.g., Henrion, 1982; Finkel and Evans, 1987; Taylor et al., 1993; James and Freeze, 1993; Dakins et

al., 1994, 1996; Brand and Small, 1995). These calculations are set in the context of an uncertain model output used for a risk-based decision, in which a loss function is specified (e.g., value of net costs minus benefits), and decisions made to minimize the expected loss (or, for optimists, to maximize expected gain). The expected value of perfect information (EVPI) can be computed for the cases where uncertainty in the input or set of inputs is completely eliminated, or for the case of imperfect information (EVSamplingI), where inaccuracies in the information gathering process for the input (e.g., data collection program, lab studies, fundamental research) are considered. This allows for explicit tradeoff between the accuracy of information gathering efforts and their costs. Programs whose data worth (EVSI) for the decision(s) exceed their cost are worth the effort. Those with the highest data worth relative to cost should receive the highest priority. Calculation of the value-of-information requires explicit specification of the quality of the information gathering program (i.e., accuracy and precision, if known), loss functions for outcomes and decisions, and the subsequent dependency of the decision on the model output and related information or data. Also, the Bayesian methods needed to assess VOI can be computationally intensive, especially for the case of imperfect information (as such, EVPI is often computed to provide a first, simpler estimate of parameter importance), though software to assist in this calculation is becoming available. In the end, decision-driven uncertainty and sensitivity analysis methods can provide the most rigorous and comprehensive basis for evaluating parameter importance; their gradual implementation, at least for high-profile/ high-cost decisions, can be expected as the methods are more broadly disseminated and applied.

**References**


1. Beck, M.B. 1987. Water quality modeling: A review of the analysis of uncertainty. Water Resour. Res. 23(8):1393-1442.

2. Brand, K.P., and M.J. Small. 1995. Updating uncertainty in an integrated risk assessment: Conceptual framework and methods. Risk Anal. 15(6):719-730.

3. Cox, D.C., and P. Baybutt. 1981. Methods for uncertainty analysis: A comparative study. Risk Anal. 1:251-258.

4. Dakins, M.E., J.E. Toll, and M.J. Small. 1994. Risk-based environmental remediation: Decision framework and role of uncertainty. Environ. Toxicol. Chem. 13(12):1907-1915.

5. Dakins, M.E., J.E. Toll, M.J. Small, and K.P. Brand. 1996. Risk-based environmental remediation: Bayesian MC/P analysis and the expected value of sample information. Risk Anal. 16(1):67-79.

6. Finkel, A.M., and J.S. Evans. 1987. Evaluating the benefits of uncertainty reduction in environmental health risk management. J. Air Pollution Control Association 37:1164-1171.

7. Henrion, M. 1982. The value of knowing how little you know: The advantages of probabilistic treatment of uncertainty in policy analysis. Ph.D. thesis, Carnegie Mellon University, Pittsburgh, PA.

8. Hornberger, G.M., and R.C. Spear. 1980. Eutrophication in Peel Inlet-2. Identification of critical uncertainties via generalized sensitivity analysis. Water Res. 14:43-49.

9. James, B.R., and R.A. Freeze. 1993. The worth of data in predicting aquitard continuity in hydrological design. Water Resour. Res. 29(7):2049-2065.

10. McRae, G.J., J.W. Tilden, and J.H. Seinfeld. 1982. Global sensitivity analysis: A computational implementation of the Fourier Amplitude Sensitivity Test (FAST). Computers and Chemical Engineering 6:15-25.

11. Merz, J.F., M.J. Small, and P.S. Fischbeck. 1992. Measuring decision sensitivity: A combined MC/P-logistic regression approach. Medical Decision Making 12(3):189-196.

12. Morgan, G.M., and M. Henrion. 1990. Uncertainty: A guide to dealing with uncertainty in quantitative risk and policy analysis. New York, NY: Cambridge University Press.

13. Taylor, A.C., J.S. Evans, and T.E. McKone. 1993. The value of animal test information in environmental control decisions. Risk Anal. 13(4):403-412.

## COMMON SAMPLING-RELATED ISSUES THAT ARISE WHEN CONDUCTING EXPOSURE ASSESSMENTS INVOLVING SOILS

Teresa Bowers

Workshop participants identified the following as common issues that arise when conducting exposure assessments involving soils (including those with Monte Carlo analysis):

- "Extent of contamination" samples are frequently the focus of site investigations and subsequently represent the data used in exposure assessment. This occurs because the initial sampling done for a site is often an attempt to define and delineate the extent of contamination. This leads to a sample set that is biased geographically and includes more sampling in contaminated areas than in clean areas. Due to cost considerations, however, this sample set may nonetheless be used. A mechanism is needed to take the bias out of such sampling plans. An area-weighted average concentration can easily be calculated, but EPA guidance requires use of an upper confidence limit on the average concentration, and it is not clear how to create an area-weighted upper confidence limit, particularly when the sample size is small.

- Often an exposure assessment is based on too few samples, or on samples taken in the wrong area. How many samples are too few? Clearly an adequate sample size is in part a function of the sample standard deviation, but it should probably also be a function of the area. Exposure areas, or exposure units, need to be chosen before the sampling plan is developed so that, with the receptor in mind, the correct areas are focused on in the sampling.

- Often many site samples have been taken, but as a site is divided into smaller exposure units the number of samples per exposure unit decreases, and the upper confidence limit, and therefore calculated risk, increases on an exposure unit basis. This puts us in the odd situation that while risks calculated for an entire site may be acceptable, the risk for each exposure unit is higher and may be unacceptable. Do neighboring samples in the next exposure unit provide information about and hence help determine confidence on the mean in any particular exposure unit? This is essentially the same question as asking on what scale does contaminant heterogeneity in soil occur.

- EPA guidance gives procedures for calculating the mean concentration's upper confidence limits when calculating risks for normal and lognormal contaminant distributions. Although there are many physical reasons that suggest that contaminants are lognormally distributed in the environment, observation suggests that distributions often deviate from lognormal in a manner that results in many low concentrations and a few very high concentrations. Is this often the case? Does it result from transport, degradation, and volatilization affecting the

distribution? Is it a result of combining distributions representing contaminated and uncontaminated areas? The H statistic method of calculating an upper confidence limit of a lognormal distribution may bias the upper confidence limit in these instances, thus biasing calculated risk. The frequency and/or significance of this bias is yet unknown.

## THE USE OF EXPERT JUDGMENT IN EXPOSURE ASSESSMENT

Max Henrion and Clark Carrington

There are a number of methods used to elicit expert judgement. All quantitative risk analysis requires some degree of expert judgment, if only in deciding how relevant a set of empirical observations is to the quantity of interest. Using expert judgment within risk analysis is unavoidable, but we can choose to make the judgment explicit or leave it implicit.

For example, consider a risk analysis to assess exposure of individuals to TXC who live near a toxic waste dump. We need to know the dispersion rate of TXC in the ground between the dump and possible exposure sites. Suppose we have estimates based on empirical measurements for several waste sites in the same region that are believed to have similar geology, but we do not have the resources to perform measurements at the site of interest. We could use a probability distribution fitted to these empirically based estimates for other sites as a representation of the uncertainty for the dispersion rate for the site of interest. Still, we cannot avoid the use of judgment, for we are implicitly judging that in terms of dispersion rate the site of interest is a random selection from the population for which we have measurements. Expert consideration of the geological characteristics of the sites may confirm that this is a reasonable judgment—but it is a judgment nonetheless.

More likely, expert consideration of the geology will conclude that the site of interest is atypical in some ways of the population of sites. A more appropriate distribution might be wider than the original empirical distribution because there is uncertainty about how the empirical distribution fits the site of interest—or it might be narrower, because we have specific information about the geology of the site of interest, suggesting that it is among the sites with high dispersion rates.

In such cases, the analyst has three choices[10]:

- Use the empirical distribution as the "objective" representation of the uncertainty about the true dispersion rate, and pretend that no judgment has been used.
- Use the empirical distribution as above, but acknowledge in the report that this is an assumption. This approach is honest, but not very useful, and does not provide the full knowledge available.
- Elicit from one or more available experts a judgmental probability distribution about the dispersion rate for the site of interest. The experts should be provided the empirical distribution of dispersion rates and all information on the geology of the sites. They may start with the empirical distribution and modify it to incorporate their judgment of how the site of interest might differ and their uncertainty about the difference.

This example situation is typical of most risk assessments: Some relevant data are known, but they are only indirectly related to the quantity of interest. Consequently, expert judgment must be used to assess the degree of relevance, the degree of adjustment appropriate for the quantity of interest, and the degree of additional uncertainty this introduces.

If the use of expert judgment is chosen, we suggest that these judgments be quantified in the form of judgmental probability distributions using careful and formal elicitation methods. There exists a well-developed and extensively practiced body of techniques to elicit expert judgment in the form of probabilities, useful for judgment by experts with no background in probability. These methods have been developed by decision analysts and risk analysts and have been in use for at least 20 years. (For descriptions and reviews of methods, see Spetzler and von Holstein,

---

[10]After the workshop, one panelist offered the following comment on the second and third choices in this list.

> "I believe that these statements are too dogmatic, and I disagree that employing empirical (i.e., surrogate) data with the explicit assumption that it is applicable to a given analysis is 'not very useful.' While such data are not completely relevant and their use thus introduces an unquantifiable element of uncertainty, the use of expert judgement to explicitly modify the surrogate data is based on subjective (albeit 'expert' judgment whose underlying assumptions and biases are not necessarily identifiable or open to quantitative analysis. Such an approach, therefore, likewise introduces another (and not necessarily less significant) source of unquantifiable uncertainty and bias."

1975, and Morgan and Henrion, 1990). Elicitation of probabilities is a demanding process with many pitfalls. Reliable results require trained elicitors following established protocols. Risk assessors should report the elicitation methods that they use. Note that the well-known Delphi approach is only one among many methods for elicitation of expert opinion, and not one that is now widely practiced, for reasons mentioned below. It is therefore unwise to use "Delphi" as a synonym for the elicitation of expert judgment.

There also exists a substantial body of experimental research on human judgment under uncertainty that demonstrates the existence of systematic cognitive biases, due to common mental heuristics—for example, overconfidence due to not considering unexpected situations. This research also demonstrates the efficacy of methods to reduce these biases. For example, asking the expert to consider extreme scenarios—"What could cause the exposure to turn out to be twice as high as the upper bound you have suggested?"—reduces the tendency be overconfident. Widely used protocols for probability elicitation include such methods for mitigating these cognitive biases.

Methods for elicitation of expert judgment vary widely in the level of effort needed and reliability of results—match the method to the needs of the analysis. Start with a "quick and dirty" method to obtain initial distributions for all uncertain variables for sensitivity analysis. Then use more expensive and elaborate methods with larger numbers of experts only where the sensitivity analysis demonstrates the importance of the quantity to the results.

It is important to use the most knowledgeable experts, with representatives of alternative credible schools of thought. Do not attempt a "random sample" from the population of experts.[11] The goal is to provide the best representation of current scientific knowledge, not to "vote on the

[11]After the workshop, one panelist offered the following comment on this statement.

> "It is naive to think that those parties who have a stake in the outcome of a risk assessment will not consciously, or unconsciously choose those experts whose view supports their desired outcome. The guidance expressed here of not attempting to choose a random sample of experts raises the potential for this sort of bias. Potentially, this can be addressed by having a disinterested third party conduct the selection, but this creates a cumbersome and time consuming logistical structure."

truth." If the disagreements among experts do not significantly affect the conclusions, it is acceptable to combine opinions, using a simple linear combination. But if there is significant disagreement among experts, it is preferable to propagate these different opinions through the analysis and to report the effects of the disagreement in the results.

Do not recalibrate expert opinion for a single study. Recalibration means expanding probability distributions to compensate for possible expert overconfidence. It is virtually impossible to assess an appropriate degree of recalibration for an individual study. Experts who know that they are to be recalibrated may resent it, or may reduce their reported uncertainty to forestall the compensation. An institution such as the EPA or SRA, however, could reasonably conduct an experimental study of degrees of overconfidence in uncertainty assessment and use the results to provided a recommended degree of recalibration for a wide class of studies.

Do not use Delphi methods that allow experts to modify their opinions after viewing the opinions of other experts. Empirical research has demonstrated that the Delphi approach often leads to extreme overconfidence due to a "group think" phenomenon. Typically, the spread of opinion among experts who have not reviewed each other's views gives a better indication of the overall uncertainty.

## METHODS FOR DEALING WITH CORRELATIONS IN MONTE CARLO ANALYSES

Christopher Frey and Scott Ferson

In general, ignoring correlations and dependencies among input variables in a Monte Carlo analysis is unacceptable. Two methods for dealing with correlations are:

- **Simulation of Correlations.** These include the restricted pairing technique (Iman and Conover, 1982; Iman and Shortencarier, 1984), Kendall's tau and Spearman's rho rank correlation (Nelsen, 1986); Pearson product moment correlation (Scheuer and Stoller, 1962); and iterative approaches (e.g., Lurie and Goldberg, 1997). The correlation matrix must be positive semidefinite (Iman and Davenport, 1982), and the software employed should check for this. The algorithms do not always yield perfect results (perhaps due to sampling error or constraints in the approach). Therefore, you should check that the correlations simulated are what were planned.

- **Dispersive Monte Carlo Sampling.** If variables are linearly related, but the strength of correlations between them is unknown (i.e., correlation coefficients are known only within intervals), then you can use dispersive Monte Carlo sampling (Bukowski et al., 1995; Ferson, 1994; cf. variance minimization described by Bratley et al., 1983) to find out how dispersed the model output might be in the general case. Whitt (1976) gives a convenient method for conducting dispersive Monte Carlo sampling. The restricted pairing technique cannot be used in this approach.

Linear or monotonic dependencies are not the only possible types of dependence among variables used in a Monte Carlo analysis. For example, a correlation of zero does not imply independence. Failing to account for nonlinear dependence can lead to a substantial over- or underestimate of variance and tail probabilities in model results. Rank correlations, although they do account for some monotonic relationships, are not flexible enough to account for all types of dependence. Gender, subpopulations, functional dependence (which should be modeled if known), and switching between alternative processes can yield complex dependency problems. It is helpful to develop mechanistic or empirical models to represent dependence between variable quantities whenever possible. In general, the analyst should try to avoid relying on simulating correlations if possible.

Other helpful methods for dealing with correlations:

- Create groups of parameters that covary, and treat each group as a new parameter (e.g., intake rate per unit bodyweight, rather than intake rate and bodyweight treated as separate variables).
- Use joint distributions and/or marginal distributions that are conditional on the values sampled from other distributions (e.g., hierarchical simulations [Voit et al., 1995]). With rich data sets, develop empirical models.
- Consider the dependence of probability distribution, shape, and parameters on other random variables.
- Stratify the population into sets of relatively homogeneous subgroups that have similar characteristics (e.g., intake rate – pica children) to reduce dependencies among variable quantities.
- Create dependency bounds. When not enough is known to either characterize dependence or to assume it does not exist, then dependency bounds analysis can be used to bound your result regardless of the underlying dependency structure (Frank et al., 1986; Williamson and Downs, 1990; Ferson and Long, 1995). This can be used in subsequent calculations using non-Monte Carlo methods.
- Use a maximum entropy approach for selecting a dependency structure subject to constraints arising from empirical evidence or compelling arguments (Yi and Bier, 1997). This is possible using copulas (Schweizer and Sklar, 1983).

In all cases, the analyst should disclose assumptions regarding correlation and dependence and give reasons for the assumptions.

**References**


1. Bratley, P., B.L. Fox, and L.E. Schrage. 1983. A guide to simulations. New York, NY: Springer-Verlag.
2. Bukowski, J.L. Korn, and D. Wartenberg. 1995. Correlated inputs in quantitative risk assessment: The effects of distributional shape. Risk Anal. 15:215-219.
3. Ferson, S. 1994. Naive Mote Carlo methods yield dangerous underestimates of tail probabilities. In: J.A. Cooper, ed. Proceedings of the High Consequence Operations Safety Symposium, Sandia National Laboratories, SAND94-2364. pp. 507-514.

4. Ferson, S., and T.F. Long. 1995. Conservative uncertainty propagation in environmental risk assessments. In: Hughes, J.S., G.R. Biddinger, and E. Mones, eds. Environmental Toxicology and Risk Assessment - Third Volume, STRM STP 1218, American Society for Testing and Materials, Philadelphia, PA.

5. Frank, M.J., R.B. Nelsen, and B. Schweizer. 1987. Best-possible bounds for the distribution of a sum—a problem of Kolmogorov. Probability theory and related fields 74:199-211.

6. Iman, R.L., and W.J. Conover. 1982. A distribution–free approach to inducing rank correlation among input variables. Communications in Statistics B11(3):311–334.

7. Iman, R.L., and J.M. Davenport. 1982. An interactive algorithm to produce a positive-definite correlation matrix from an approximate correlation matrix (with a program users' guide). Sandia National Laboratories, Albuquerque, NM, SAND81-1376.

8. Iman, R.L., and J.M. Shortencarier. 1984. A Fortran 77 program and user's guide for the generation of Latin hypercube and random samples for use with computer models. AND83–2365. Sandia National Laboratory, Albuquerque, NM. January.

9. Lurie, P.M., and M. Goldberg. 1997. An approximate method for sampling correlated random variables from partially specified distributions. Management Science. In press.

10. Nelsen, R.B. 1986. Properties of a one-parameter family of bivariate distributions with specified marginals. Communications in Statistics (Theory and Methods) A15:3277-3285.

11. Scheuer, E.M., and D.S. Stoller. 1962. On the generation of normal random vectors. Technometrics 4:278-281.

12. Schweizer, B., and A. Sklar. 1983. Probabilistic metric spaces. New York, NY: North-Holland.

13. Voit, E.O., W.L. Balthis, and R.A. Holster. 1995. Hierarchical MC/P modeling with s-distributions: Concepts and illustrative analysis of mercury contamination in king mackerel. Environ. Intl. 21(5):627–635.

14. Whitt, W. 1976. Bivariate distributions with given marginals. The Annals of Statistics 4:1280-1289.

15. Williamson, R.C., and T. Downs. 1990. Probabilistic arithmetic I: Numerical methods for calculating convolutions and dependency bounds. Intl. J. Approximate Reasoning 4:89-158.

16. Yi, W., and V. Bier. 1997. An application of copulas to accident precursor analysis. Management Science. In press.

## APPROACHES TO ENSURING THE STABILITY OF MONTE CARLO RESULTS AT THE TAILS

Workgroup Chair: David Burmaster

There are two aspects of the issue regarding the uncertainty and/or stability of Monte Carlo results at the tails: measured data and simulation. While it is inevitable and universally true that we know much more about the center of a distribution than its tails, we need only to take suitable precautions to ensure the integrity of the results for management purposes.

### Measured Data

The following discussion presumes that data have been collected using an appropriate random or stratified-random sampling design.

With N data points measured for a phenomenon, we have no hard information above the empirical {1–0.5/N}th percentile of the variability in the phenomenon, except insofar as we are willing and able to model the variability with a parametric distribution that extrapolates beyond the range of measurements. For N (less than or equal to) 10, for example, this is a major limitation for both deterministic and probabilistic studies. In this range, however, probabilisitic methods have a distinct advantage over deterministic approaches, especially if the analyst distinguishes between variables with respect to variability versus uncertainty as discussed in Sections 3.3.1 and 3.3.2. For N = 20, we have empirical information spanning the central 95 percent of the variability in the distribution. As N increases and the analyst has more confidence in the distribution, the uncertainty represented by the random parameters decreases (i.e., the distributions for the parameters converge towards point values). In other words, as N increases, the uncertainty in the distribution for variability in the data decreases. The uncertainty never disappears for finite N, however, and the uncertainty decreases most slowly in the tails of the distribution for variability. So it is not surprising that measured data with large N are always better in probabilistic (and deterministic) studies.

**Simulation**

With simulation, the analyst can examine and visualize the implications of the "uncertain" variables chosen as inputs. For the first few realizations (or iterations) of the simulations, the uncertainty in the output distributions—especially in the tails—is dominated by the randomness in the simulation. As the number of realizations grows, the uncertainty in the output distribution decreases asymptotically to the combined uncertainties inherent in the input distributions. In this case, the analyst can use numerical experiments to demonstrate acceptable convergence in both performing i) the inner-loop for variability and ii) the outer-loop for uncertainty employing a 2-D Monte Carlo analytical scheme (see "Use of Numerical Experiments in Monte Carlo Analysis" earlier in this Appendix). Performing these numerical experiments is inexpensive with the computers and software currently available. (See Section 3.2.1 for more on numerical experiments.)

## ROLE OF BAYESIAN METHODS IN MONTE CARLO ANALYSES

Workgroup Chair: Mitchell Small

Controversy persists over the extent to which probability distributions are appropriately used to represent uncertainty in expert knowledge versus their more traditional use in fitting data. While classical statistical methods for fitting distributions attempt to consider only the information contained in the data, Bayesian statistical methods explicitly allow for the incorporation of subjective expert knowledge and judgment when developing distributions. Because they allow the knowledge in the expert judgment to be combined with the information in the data, however, Bayesian methods have the capacity to bridge the gap between those who focus on expert knowledge in developing distributions and those who put greater emphasis on lab or field data.

With Bayesian methods, a prior distribution is assumed or elicited to represent expert judgment about the distribution before seeing the data. The prior distribution can incorporate information from previous studies, the scientific literature, or data from other study sites. Alternatively, lacking any such prior information, various forms of "informationless" priors can be assumed. A likelihood function is next identified to relate the probability of obtaining different study (e.g., data) outcomes given each possible value of the prior distribution. Once data are obtained, these are combined with the prior distribution and the likelihood function to obtain the posterior distribution—the estimated distribution which combines the information in the data with that of the (prior) expert judgment. General texts on Bayesian methods are available (DeGroot, 1986; Press, 1989; Berry, 1996), as are edited volumes with applications to environmental quality and health (Gatsonis et al., 1993; Berry and Stangl, 1996).

Bayesian methods are quite compatible with efforts to separate variability and uncertainty in exposure and risk assessment. A prior distribution can be used to describe the uncertainty in the parameters of a variability distribution; then, as data are collected, a posterior distribution of the parameters is obtained (generally, with less uncertainty). Examples of this include Wood and Rodriguez-Iturbe (1975), Iman and Hora (1989) and Small (1994). Methods for eliciting prior distributions for uncertain variability distributions are presented in the statistics literature (Kadane et al., 1980; Chaloner, 1996; Wolfson, 1995).

Bayesian methods are very useful for designing experiments and data collection programs to reduce uncertainty. This can be accomplished using the value-of-information approach discussed in A Hierarchy of Methods for Sensitivity Analysis (see pp. E-10 - E-15). In addition, when different stakeholders have very different prior beliefs about a particular exposure distribution, Bayesian methods can help illustrate the quantity and quality of data required so that both sides will have essentially the same posterior distribution despite their differing priors (Wolfson et al., 1996). When data sets are large and/or with sufficiently accurate data as defined by the likelihood function, the posterior distribution is relatively unaffected by the assumed prior. Thus, Bayesian methods exhibit a shift from reliance on the subjective expert opinion to reliance on the data, as more and better data are collected.

Bayesian methods can be computationally intensive. For certain types of distributions, the mathematical form of the prior distribution is maintained as data are collected, so that the mathematical form of the posterior is the same as that of the prior. In this case the prior and posterior distributions are said to be "conjugate." In the more general case, numerical methods are required to compute posterior distributions. In either case, Bayesian methods are conceptually difficult to grasp for many, and exposure and risk assessors will require the assistance of a competent statistician experienced with Bayesian methods to apply these techniques to their applications. While this may preclude their application in many cases, increased use of Bayesian methods is likely in the future because of the types of problems they can address and the associated insights and benefits they provide.

**References**


1. Berry, D.A. 1996. Basic Statistics: A Bayesian Perspective. Belmont, CA: Duxbury Press.

2. Berry, D.A., and D.K. Stangl, eds. 1996. Bayesian Biostatistics. New York, NY: Marcel Dekker, Inc.

3. Chaloner, K. 1996. Elicitation of prior distributions. In: Berry, D.A., and D.K. Stangl, eds. Bayesian Biostatistics. New York, NY: Marcel Dekker, Inc. pp. 141-156.

4. Degroot, M.H. 1986. Probability and Statistics, Second Edition. Reading, MA: Addison-Wesley.

5. Gatsonis, C., J.S. Hodges, R.E. Kass, and N.E. Singpurwalla, eds. 1993. Case Studies in Bayesian Statistics. New York, NY: Springer-Verlag.

6. Iman, R.L., and S.C. Hora. 1989. Bayesian methods for modeling recovery times with an application to the loss of off-site power at nuclear power plants. Risk Anal. 9(1):25-36.

7. Kadane, J.B., J.M. Dickey, R.L. Winkler, W.S. Smith, and S.C. Peters. 1980. Interactive elicitation of opinion for a normal linear model. J. American Statistical Association 75:845-854.

8. Press, S.J. 1989. Bayesian Statistics. New York, NY: John Wiley & Sons, Inc.

9. Small, M.J. 1994. Invariably uncertain about variability? Try the normal-gamma conjugate. In: Proceedings of the 87th Annual Meeting of Air & Waste Management Association, Air and Waste Management Association, Pittsburgh, PA. Paper 94-TP55.05.

10. Wolfson, L.A. 1995. Elicitation of priors and utilities for Bayesian analysis. Ph.D. thesis, Carnegie Mellon University, Pittsburgh, PA.

11. Wolfson, L.A., J.B. Kadane, and M.J. Small. 1996. Expected utility as a policy-making tool: An environmental health example. In: D.A. Berry, and D.K. Stengl, eds. Bayesian Biostatistics. New York, NY: Marcel Dekker, Inc. pp. 261-277.

12. Wood, E.F., and I. Rodriguez-Iturbo. 1975. Bayesian inference and decision making for extreme hydrologic events. Water Resour. Res. 11(4):533-542.

## RECOMMENDATIONS FOR PRESENTING MONTE CARLO RESULTS TO RISK MANAGERS

Summarized from Bloom et al. (1993) in Appendix F.

**At the beginning of a briefing:**

- Present an overview of the significance of the analysis.
- Identify stakeholders and briefly describe who is saying what about this issue.
- Discuss the positions of other EPA offices and other important constituents on this issue.

**When characterizing the risk of a chemical:**

- Present information concerning the severity of the adverse health effect posed by the chemical (i.e., is it death or is it a runny nose?).
- Establish the extent to which scientists believe the chemical is really a hazard to humans.
- Describe the level of confidence in the data and in the numerical assessment of risk. (Tell who else has seen the information and who else agrees with it.)
- Explain where the data gaps are and tell how important those gaps are to the overall risk estimate.
- Highlight potential "high visibility data gaps" that are likely to become the focus of attention of groups outside EPA.
- Show all the formulae used to estimate exposure point concentrations, exposure doses, toxic potencies, hazard indices, and/or incremental lifetime cancer risks. As for any risk assessment, show the formulae and the spreadsheets in the text, in tables, or in an appendix.
- Calculate and present the point estimates of exposure and risk that are generated following the current deterministic risk assessment guidelines from the appropriate regulatory agency. The calculation of point estimates using standard techniques is a desirable first step in undertaking a probabilistic risk assessment.

**When discussing exposure:**

- Define what population is at risk. (Is it the general population? Children? The elderly? Minority groups?)
- Estimate the number of people who are exposed to levels of concern and present the range of uncertainty around the exposure numbers.
- For parameter uncertainties, what distributions are used and why?
- What models have been used, what assumptions do they make, and what uncertainty do they possess? Clarity in communicating model assumptions and structure is as important as clarity in communicating uncertainty about results.
- What issues and factors do the models include?
- Are the model equations implemented correctly?

**When presenting risk management options:**

- Present the legislative mandate.
- Identify potential risk management options, including ones that have already been rejected along with reasons for rejection.
- Discuss each option in regard to its costs, ease of implementation, and likelihood of success for reducing risk.
- Clarify how much each option will reduce risk rather than merely shift it from one medium to another.
- Estimate what proportion of the risk from this chemical a particular action will actually address.
- Discuss the consequences of doing nothing.
- Review what has been done in previous similar situations.
- Mention studies in progress which could yield new and important information about this chemical.

**In informing the decision-maker about the weaknesses of data:**

- Provide a sense of the uncertainties of the data.

- Present the results from univariate (or multivariate) sensitivity analyses of the deterministic calculations to identify the inputs suitable for probabilistic treatment and then discuss any variables not included in the sensitivity analysis. A typical risk assessment may require the specification of over 100 input variables. Only a few of these inputs drive the risk assessment in one or both of these senses: i) The values of some inputs account for a dominant fraction of the predicted risks and/or ii) the ranges of some inputs account for a dominant fraction of the range in the predicted risks. When using probabilistic techniques, it is important to understand which inputs drive the predicted risk in both of these senses.

- Discuss how effective any proposed alternative substitutes to this chemical would be in accomplishing their intended purpose.

- Discuss any risks, trade-offs, or unintended consequences of the proposed substitutes to this chemical.

- Mention decisions (especially conflicting decisions) that other EPA offices have made about this chemical.

**When presenting risk information:**

- Charts that are complex and visually busy contain more detail than is necessary to make decisions. For example, box plots are thought to convey too much detail for risk managers' purposes. When presenting the results of a Monte Carlo analysis, most managers prefer the cumulative distribution format. Large spreadsheets and proprietary computer codes can present obstacles to clarity.

- Since different people process information in a variety of ways, it is appropriate for risk assessors to provide more than one format for presenting the same information. In general, visual presentation aids need to be straightforward and provide information in a clear, uncluttered way. Too much information in one chart tends to drown out the most important point. Charts, graphs, and other visual aids should be accompanied by written information and/or oral briefings. No one expects a chart to convey all the important information at a glance.

- When it is necessary to provide a more complex visual, it may be necessary to provide a written report—including tables and figures—before a briefing, so that interested constituents can study the information, absorb the details, and consider the questions they want to ask.

## RECOMMENDATIONS FOR PRESENTING INFORMATION ABOUT INPUT DISTRIBUTIONS

Summarized from Burmaster and Anderson (1994) and Henrion (Appendix D).

- Provide the following information for each input distribution:

  - A graph showing the full distribution and the location of the point value used in the deterministic risk assessment.

  - A table showing the mean, the standard deviation, the minimum (if one exists), the 5th percentile, the median, the 95th percentile, and the maximum (if one exists).

- Include a 5- to 10-page justification of the selected distribution based on results in a refereed publication, from new developments, or from elicitation of expert judgement in the risk assessment.

- Discuss how, for parametric distributions, the statistical process or the physical, chemical, or biological mechanism creating the random variable influences the choice of the distribution.

- Show, to the extent possible, how the input distributions (and their parameters) capture and represent both the *variability* and the *uncertainty* in the input variables.

- Discuss the methods and report the goodness-of-fit statistics for any parametric distributions for input variables that were fit quantitatively to measured data. Show plots of the parametric fits and the data on the same axes. Discuss the implications of any important differences. If any distribution was generated qualitatively or by expert judgement, discuss the techniques used.

- Discuss the presence or absence of moderate-to-strong correlations between or among the input variables.

- Present the name and the statistical quality of the random number generator used. Some well-known commercial products have inadequate random number generators with short recurrence periods. If the analyst writes his or her own specialty generator, include an appendix in the report listing the algorithm and the implementation, along with the results from a quality assurance audit.

- Discuss the limitations of the methods and of the interpretation of the results. Be sure to acknowledge the source, the nature, and the possible effects of any

unresolved sources of bias not explicitly included in the analysis, and indicate where additional research or measurements could improve the analysis.

**Reference**

1. Burmaster, D.E., and P.D. Anderson. 1994. Principles of good practice for the use of Monte Carlo techniques in human health and ecological risk assessments. Risk Anal. 14(4):477-481.

## DISTINGUISHING A "GOOD" FROM A "BAD" MONTE CARLO ANALYSIS

Workgroup Chair: Dale Hattis

### Objectives and Purpose

- Determine if the Monte Carlo analysis is consistent with and appropriate for the stated purpose of the exposure assessment. Check to determine if the Monte Carlo analysis includes a clear statement of purpose and objectives.

### Checks on Input Distributions and Assumptions

- Make sure the analysis enables you to determine how the distributions were derived or obtained from referenced sources. Carefully review the bases for any site-specific or novel distributions.
- Check the variability and uncertainty of specific parameters against those reported in other studies and analyses.
- Check the bases of truncated distributions for the truncation and determine whether values have been allowed to take on physically impossible values.
- Check to determine if the authors have assessed the effects of clear dependencies among parameters. Where these have been identified, determine if they have been included in the exposure models in a logical manner.
- Check to determine if the underlying assumptions have been appropriately identified.
- Check on the quality of information available for the tails of the input distributions. These will affect the tails of the output distributions. Examine if this has been specifically considered by the analysts.
- Check to see if the analysts has used influence diagrams, numerical experiments and/or sensitivity analysis in an appropriate manner to evaluate the relative importance of different parameters and exposure pathways in the analysis.

### Checks on Model and Computational Mechanics

- Examine the boundaries of the modeled system with respect to the stated purpose and objectives of the exposure assessment. For example, have the populations of decision-making interest been reasonably defined? Or have there been geographic, age, or exposure pathway truncations or aggregations that could materially change the implications of the results for decision-making?

- Check for model and analytical transparency. Make sure that all equations and distributional assumptions are clearly documented.

- In exposure assessments for multiple pathways where the same parameter appears more than once, check to determine that the same value of that parameter is used consistently on each trial. This requires being able to understand how the Monte Carlo analysis was implemented. The report should provide sufficient information for the reviewer to judge this aspect of the analysis.

**Checks on Results**

- Check to make sure that variables and expressions balance dimensionally.

- For pure multiplicative/division models, perform a calculation of exposure using the median or most probable values of the distributions in a separate deterministic equation. The result should be similar—but not necessarily identical—to the median value produced by the Monte Carlo analysis.

- Check to make sure that the temporal aspects and units for the exposures and doses are consistent with those used to represent the effects in risk calculations.

- Perform a mass balance calculation to check the exposure estimates or review checks made by the analysts, if possible.

- Check to see if the analysis includes appropriate deterministic calculations for standard high end exposure estimates.

- Check to determine if the analyst has adequately distinguished between variability and uncertainty at a conceptual and possibly at a quantitative level as appropriate for the analysis.

- Check to determine if the analysts has identified the exposure parameters and pathways that are most important with regard to exposure estimates and the uncertainty associated with these estimates.

- Check to be sure that the report includes a clear statement of limitations, unaccounted for uncertainties, and possible biases.

- Check to determine if and how the stability of results at the tails of the distribution have been evaluated. This is critical for the needs of risk managers.

**Other Checks**

- Check if the results or values used in the analysis are given in more significant figures than warranted. This may indicate a certain lack of sensitivity and a certain desire on the part of the analyst to retain data that are not really meaningful.

**Extra References**

California Department of Health Services. 1991. Health risk assessment of aerial application of malathion-bait. Pesticides and Environmental Toxicology Section. Berkeley, CA.

Marty, M.A., S.V. Dawson, M.A. Bradman, M.E. Harnly, and M.J. DiBartolomeis. 1994. Assessment of exposure to malathion and malaoxon due to aerial application over urban areas of Southern California. J. Exposure Analysis and Environmental Epidemiology 4:65-81.

## APPENDIX F

## REFERENCES

# COMMUNICATING RISK TO SENIOR EPA POLICY MAKERS: A FOCUS GROUP STUDY

Produced for:

Office of Air Quality Planning and Standards
U.S. Environmental Protection Agency

Produced by:

Diane L. Bloom, Ph.D.
Bloom Research

Dianne M. Byrne
Chief, Program Analysis and Technology Section
Emission Standards Division
OAQPS

Julie M. Andresen
Emission Standards Division
OAQPS

March, 1993

## Acknowledgments

We would like to thank all the people who helped make the EPA risk management decision making focus group a success. We appreciate the help of Bob Kellam, John Vandenberg, and Harvey Richmond for providing us with suggestions for graphics depicting uncertainty, writing scripts, and acting in the risk assessment mock-briefing video tape used in the focus group. We would also like to thank Carol Jones for her assistance and technical support, and Jonathan Bloom for filming and editing. We would further like to acknowledge all of the people from OAQPS who reviewed the topic guide and made suggestions for the focus group. Finally, we would like to thank the EPA decision makers who took time out of their busy schedules to participate in the focus group and share their ideas and insights with us.

# Table of Contents

# Communicating Risk to Senior EPA Policy Makers: A Focus Group Study

## 1. INTRODUCTION

Information needs of risk managers in high-level EPA positions are not always well understood by risk assessors who provide them with risk assessment results. Little is documented about how risk managers use risk information in making regulatory decisions, or which formats they find most useful. Moreover, not enough is known about the level of detail risk managers want about the assumptions and uncertainties of the risk data.

The Office of Air Quality Planning and Standards (OAQPS) of the U.S. Environmental Protection Agency is charged with presenting risk assessment information to decision makers. In the past, the office has not asked decision makers about how well this information serves their needs. To enable risk assessors to present the essential information needed by risk managers, officials within OAQPS authorized a focus group to ask risk managers directly about their needs and preferences. The focus group was not designed to test theories or preconceived ideas about how risk information should be presented. Instead, it was a search for the ideas and insights of those who use this information everyday for real-world risk management decisions.

The focus group consisted of eleven senior EPA Headquarters decision makers, selected to represent the following offices and programs throughout the Agency: the Office of Air and Radiation; the Office of Water; the Office of Solid Waste and Emergency Response; the Office of Prevention, Pesticides and Toxic Substances; the Office of Policy Planning and Evaluation; and the Office of Air Quality Planning and Standards. The participants included Assistant Administrators, Deputy Assistant Administrators, Office Directors, and Science Advisors. The focus group provided specific examples of visual risk information for the group's on-the-spot reactions, and lasted two hours. The focus group took place two months after an OAQPS in-depth telephone interview study of 30 risk managers from Headquarters and from Regional, State, and local air pollution control offices addressing similar questions.

## 2. SUMMARY OF RESULTS

### 2.1 Key issues for risk managers making regulatory decisions.

Participants were asked to imagine that they had to make a policy recommendation or decision about regulating a hazardous pollutant such as benzene, asbestos, or chlordane. They were asked which major issues they would consider in making their decision. The issues they think about fall into several categories, including legal considerations, adverse health effects, exposure information, possible risk management options, degree of consensus, and issues of confidence in the data. One further issue they said they consider is what precedents they may be setting with this decision.

#### 2.1.1 Legal considerations

Knowing the legislative mandate is an important first step according to most of the participants, as illustrated by this comment:

> "The legislative mandate is critical. For example, is it a risk/benefits statute or risk only? You regulate polluters differently under different statutes. We would regulate radon under the Indoor Radon Abatement Act very differently than we would under the Safe Drinking Water Act."

#### 2.1.2 Adverse effects and exposure issues

Other issues the policy makers consider are the magnitude of the adverse effect of the chemical and the extent of exposure to the population. They seek a sense of where this chemical fits in the spectrum of environmental problems. Participants described these issues in the following way:

> "I want to know what are the physical effects and how serious are they. Is it death or is it a runny nose? How many people are exposed and at what levels?"

> "I want to know who is exposed. Is it the general population? Children? The elderly? Are there sensitive subpopulations? Are certain groups like the poor or ethnic minorities exposed differently?"

2.1.3 Risk Management Options

The group members said they also look at several aspects of the risk management options that are available. They look at the costs of those options and the benefits in reduced risks. In addition, they consider the feasibility of implementation of each option. The following comment illustrates some of the issues surrounding risk management options:

> "I want to know about all the risk management options including those that have already been rejected. I want to know how successful the risk management options are likely to be in reducing risk. Then, I consider our ability to implement those options. Sometimes we have good options but we can't implement them. How many resources will it take?"

Along these lines, participants also said they consider whether there are substitutes or alternatives available for the chemical they are considering regulating. Ideally, they would like information regarding the efficacy and safety of those alternatives.

2.1.4 Degree of Consensus

Participants said that before making their decision, they identify who it is that cares about this issue. Sometimes, if a lot of vocal people are very concerned about the risks of a chemical, EPA addresses those concerns. One risk manager offered this comment:

> "Is this an issue of public concern? Who cares about the risks? We're spending a lot of time on "killer carpets" where we don't have a lot of risk, however there is a lot of concern about the risk on the part of some very vocal people. So we at EPA address those concerns."

Similarly, risk managers consider what others both within the Agency and outside are saying about the issue. They consider the positions of environmental groups, citizens groups, and industry, as well as the positions of other EPA offices and the Office of Management and Budget. A typical comment follows:

> "It is important for me to know what the reactions of the stakeholders will be to the various risk management options and to our recommended options....In terms of the Agency, I think about whether the Office of Policy Planning, and Evaluation (OPPE) or the Office of General Counsel (OGC) would agree with our

recommendations. Could I sell this through the management chain? Could I get it through OMB?"

2.1.5 Confidence in the data

Another issue risk managers say they think about when making decisions is the extent to which the risk and economics information has been peer reviewed. Peer review by credible groups gives them more confidence in the information. As one participant explained:

> "What is your level of confidence in the risk information and in the numerical assessment of risk? Who has looked at this information? Just our staff or others? I want to know to what extent the science has been peer reviewed by the Science Advisory Board."

In summary, the focus group participants said that they consider a variety of issues when making regulatory decisions including:

- The legislative mandate
- The seriousness of the adverse effects of the chemical
- The number of people exposed to the chemical
- The subgroups are exposed (Children? Ethnic groups? The elderly?)
- The costs, cost effectiveness, economic impact, and ease of implementation of the risk management options
- The alternatives and substitutes available for the chemical in question
- The efficacy and safety of those substitutes
- The people who care about this issue
- The positions of stakeholders concerning the issue and their likely reactions to options
- The positions of other offices within EPA
- Likely reactions of OMB
- The extent to which the science and economics have been peer reviewed
- The precedents that are being set by this decision

**2.2 What information risk managers want to hear when they are briefed.** The participants were told to imagine that they were going to be briefed on a hazardous air pollutant. The purpose of the briefing would be to provide them with information for making a regulatory decision. They were asked to further assume that this would be the first time they would

hear the information. The scenario suggested that a large number of people were exposed to this chemical and the costs of controls to industry were high. They were told to pretend that their one-hour briefing had been cut to 15 minutes. They were then asked what key points they would like addressed within this time constraint.

All the group members noted that, in reality, they wouldn't make a decision based on a 15-minute briefing. In contrast, their decisions are complex and would be thought out more carefully, as these comments reflect:

> "When you think about making a decision, it's on the basis of information on several different dimensions. In my approach, I couldn't make the decision on just one. If I just knew the hazard and the exposure it would get my attention, but I can't move forward on my decision making without some of the other pieces of information."

> "If it is an important decision, we're going to take the time until we get a level of satisfaction with the data and information that we need to make that decision. I don't think we are going to say that X is more important than Y. We're going to want to know both X and Y. We'll get the information we need to make the decision."

The participants said that they would want quantitative assessments of risk such as the potency of the hazard, maximum individual risk, and population risk. In regard to the qualitative information they want to know in a briefing, there was some overlap with the issues they consider in making decisions (listed in the preceding section). For example, they would want to determine how many people are being exposed to the chemical and the implications of this for health and safety. Moreover, they want to have an idea of who else agrees with the scientific information.

In a briefing, the risk managers want a picture of the possible options including the option of taking no action. They said it is also helpful for them to know which options have been used in previous, similar situations. These comments reflect some of the information needs of the group:

> "What are your risk management options and how effective are they? In this situation the natural question to ask is 'What is the consequence of doing nothing?' and 'What has been done in previous other situations with similar pollutants?'"

"I want to hear about the reactions of the stakeholders to the risk management options and the recommended options. How will they react if we do nothing?"

To summarize, decision makers want to be briefed about the following kinds of information:

- Numerical estimates of risk
- Magnitude of the adverse effect
- Level of exposure
- Level of confidence in the data
- Risk management options
- Consequences of doing nothing
- Reactions of stakeholders to the recommended risk management options
- Costs and economic impacts
- What has been done in previous, similar situations
- Positions of other EPA offices and of OMB

The participants said that they would use this same information to brief the Administrator. Additionally, they would tell the Administrator what their recommendation is, and how they think that decision will be accepted by the public.

**2.3 What risk managers would like to see in a qualitative description of the risks of a pollutant.** The participants were asked to brainstorm to create an introductory statement to a briefing package which would present "the big picture" about the risks of a chemical. They were told that the purpose of this exercise was not to see whether the details are accurate, but instead to find out what kind of language they would use. The group chose benzene to use as an example.

They would describe it in the following way:

"Benzene is a known human carcinogen. A large number of people are exposed to levels of concern. Regulation will make a difference -- there is something we can do about it. We can reduce exposure on a technological basis. Controls are reasonable. Additionally, we are under a court order so we MUST do something about it."

In addition, they would include a discussion of the efficiency and cost effectiveness of controls and the reactions of industry, the environmentalists, and the Administration.

In such an introductory statement, they would also include other possible options besides the proposed option, including options that were rejected and reasons for their rejection. One risk manager indicated that some other background information might be important to clarify in an introductory statement:

> "There is a lot known about benzene, so saying it is a known human carcinogen is very meaningful here."

This participant noted that "Known" is a code word for the level of research that has been done in assigning this carcinogen a classification. In a briefing, he said it would be helpful to explain:

> "What groups say it's a carcinogen and what are the range of options that have been expressed. A concise discussion of the science would be helpful."

There was consensus that such an introductory qualitative statement, such as the description of benzene, would be useful in a briefing.

**2.4 What kinds of uncertainty information risk managers want.** There has been much discussion at EPA about the presentation of uncertainty information relevant to the risk data. The risk guidance memo (February, 1992), for example, advises risk assessors to provide risk managers with an understanding of the uncertainties of the underlying analysis. However, it is not clear to risk assessors exactly how risk managers define uncertainty, and what level of detail is optimal.

Focus group participants were asked to discuss how they define "uncertainty" in the context of risk. They cited several dimensions of this term including uncertainty around the numerical assessment of risk, the uncertainty about the magnitude of the hazard in humans, the uncertainty around the extent of the exposure to humans, the data gaps, and the uncertainty surrounding the effectiveness of the options. These dimensions are discussed below.

### 2.4.1 Uncertainty of the adverse effect

Risk managers want to know about the uncertainty surrounding the actual adverse effect of the chemical. Perhaps there is an adverse effect documented in animals, but how certain is it that humans will experience that effect? The following comments illustrate uncertainty in this context:

> "How certain are we about the hazard information? How confident are we that this is going to be a problem in humans? Benzene is in a category of a small number of compounds where we have data that are good. How confident are we that another chemical will cause cancer (or reproductive effects or birth defects) in humans? At what exposure level can we expect to see an effect in humans? How sure are we of that?"

> "Is there consensus among the scientific community about the seriousness of this effect, or are we the only ones saying this?"

The following comment illustrates how uncertainty could exist about an effect at a particular exposure level which may be well-documented in animal studies but not in humans.

> "In formaldehyde, for example, we have very good information about carcinogencity in animals. We have some information for humans. Based on our experience, we believe that it will be a human carcinogen. But, what about exposure levels? It may be that you're going to need higher levels in humans, than what you get in animals. So, it has to do with extrapolating from an effect in animals to humans. Also we must extrapolate the dose levels at which humans are going to be exposed. Are these reasonable assumptions?"

### 2.4.2 Uncertainty surrounding exposure

Risk managers also use "uncertainty" in the context of exposure. They want to know the range of uncertainty of the actual number of people who are exposed to the chemical. They seek answers to questions such as:

> "What are the bands around the numbers? When you say there's 100,000 people exposed do you really mean there may be 1,000 or one million and we picked 100,000 because it's somewhere between?"

"Do we know if anybody is exposed? Are people really being exposed to these levels?"

2.4.3 Data gaps

Another dimension of uncertainties is "data gaps." The risk managers want to know just where the data gaps are and how significant they are to the overall estimate. One participant gave the following example of a data gap:

> "There are some things that we know and feel reasonably confident about. But we know that there is missing information. We've made estimates about something, but there's a hole in the estimate. We want to identify the fact that you don't have information from humans, for example. An example of a data gap would be, 'It's all based on animals, there's no information from humans.'"

Although there will always be data gaps, they want to know, especially, about the ones that are likely to be very important. It will be useful for decision makers to know about these data gaps because they can then see what kinds of criticism they will need to confront. This comment illustrates the importance of knowing the data gaps that others outside of EPA are likely to notice as deficient:

> "What's going to pop up and bite us? Is there a missing element in the data? Data gaps are always present, but which ones are going to become the focus of attention. We want to know if some professor at Harvard is likely to go on television on the evening news and say, 'Look what EPA didn't do...'"

In regard to data gaps, risk managers want to know if new information is on the horizon. They want to know about any forthcoming new and significant studies, and whether it is worthwhile to wait for that information before making their decisions. Two comments on this issue were:

> "If we delay this decision in order to get more information, would it be worthwhile? You may have the option of saying, 'Let's get another study done'....Is there a new study coming out next year that people think will be very significant? Will that study make a difference or will it be just one more study in a chain of studies?"

> "Timing is critical. Is a new study already in the works? If they are just starting and it will be a 5 or 6 year process it would not be as useful to wait."

Along these lines, risk managers want to know which direction the risks would probably go if the data gaps were filled in. As one put it:

> "Do the uncertainties tilt in a particular direction? Have we made assumptions that are conservative or not? Are the unknowns things that have been ignored so that we may be missing some important risks? If we could fill in the uncertainties, which direction do we think the risk assessment would be going?"

2.4.4 Uncertainty surrounding management options

Another area of uncertainty that risk managers identify as important is uncertainty surrounding the options. For example, they want to know how likely the options are going to work to reduce risks. They also want to know about the uncertainties about the costs of the risk management options.

In summary, EPA policy makers have several meanings in mind when they speak of "uncertainties." Generally, they want the following questions surrounding uncertainties addressed in briefings:

- How confident are we that this chemical is really a problem?
- How confident are we that it will really cause cancer (or other effects) in humans?
- Do we really know how many people are exposed?
- What are the error bands around the exposure numbers?
- If we say 100,000 people are exposed, is the range really between 100 and 1 million?
- How accurate are the extrapolations from animal to man or from high to low doses?
- What are the data gaps, and how important are they?
- If we had more information, in which direction would the risks go?
- If we delay the decision to get more information, what would happen?
- What new studies are in progress?
- Who has reviewed the uncertainty analysis -- our staff only, the Science Advisory Board, or other credible groups?
- What are the uncertainties around costs?

A few risk managers provided examples of uncertainty information that have proven to be very useful. One piece of an effective presentation of uncertainty is described below:

> "In Pesticides, we have run sensitivity analyses around the different use scenarios. If 100% of apples are treated with pesticide, (or 50% or 10%), what does that do to your whole exposure number? In addition, we also look at different risk management options. For example, what happens if you put worker protection requirements in place to get risk down, rather than take the chemical off the market. These analyses give you both the sensitivity around your exposure numbers and the sensitivity around the success of your options."

According to participants, sometimes uncertainty information is too vague or inconclusive to be meaningful. For example, participants said it is not helpful when they are told, "Here is the plausible upperbound. The real risk could be less and might be zero." Equally unhelpful are comments they sometimes receive such as, "We can't do anything with this data, it's too uncertain," or "You should do something very stringent because of all the uncertainties."

**2.5 What additional information risk managers seek about the weaknesses of the data.** In their decision making, the focus group participants all said they want to know whether their positions are vulnerable. The risk managers want to know 'What is industry thinking? What are environmental groups thinking?' Knowing the weaknesses of data, highlights potential challenges and helps risk managers prepare arguments to support the EPA decision.

When these decision makers talk about "weaknesses in the data," they are referring to a variety of possible problems with the supporting data that would leave the policy decision open to attack by stakeholders, other EPA offices, or OMB. For example, they want an idea of where the data gaps are and how important those gaps are to the overall risk estimate. They also want to know how much an option will actually reduce risk, rather than merely shift it from one medium to another (e.g., air to water).

An important factor before making their decision could center around the substitutes available to the chemical in question. Facts about the efficacy of the proposed alternatives to the chemical they are regulating are needed. Participants said that having this type of information has helped them in the past. For example, in the pesticide area, farmers want to know how well a substitute chemical will work.

They noted with surprise that seldom does anyone ask how risky those alternatives are. For example, usually no one questions the flammability of substitute chemicals. One participant remarked:

> "Are the substitutes to this chemical risky? We didn't used to ask that. The farmers ask, 'Are the substitutes effective?' but the environmental community doesn't ask if there are risks to these alternatives? They are often more concerned with getting the chemical in question off the market. We at EPA now ask that question, but it's not typically asked by stakeholders."

Nonetheless, the risk managers believed it was important to know about any unintended consequences, tradeoffs, or risks of the proposed alternatives. This type of information would help them consider the broad implications of policy decisions so they could make better decisions, as this comment illustrates:

> "Risks of alternatives or substitutes is something we think EPA should be thinking about but sometimes we don't. We should be thinking of safety issues or tradeoffs. With asbestos, for example, we need to think about the safety of brakes using alternative substances. It's a tradeoff issue that we need to consider. If we don't look at the big picture, we could be saying 'You may die in a car accident, but that's not our fault. We saved you from dying of cancer by taking asbestos out of your brakes.'"

Several focus group members said that they often do not receive information about conflicting positions of other EPA offices. A conflict about what to do with a chemical may arise, because different offices tend to work independently and thus are unaware of other recommendations or decisions. Generally, they said that no one looks at consistency among offices in how they treat risks. These risk managers thought this lack of coordination among EPA offices could embarrass decision makers asked to explain such a disparity. Risk managers need more information about other EPA office actions on particular chemicals, as the following comment illustrates:

> "We need to be more aware of consistency among offices in terms of how we treat risks. As one office goes forward and announces a decision, it is at times uncomfortable to have someone point out that another office has either accepted that risk and felt comfortable with it, has rejected that risk and felt comfortable with that decision, or

has been very concerned about the alternatives you are now advocating....Perhaps we didn't ask the questions we should have asked, or maybe our statutes have taken us in different directions. Or maybe we have some real life circumstances that force one office to accept a risk that another office rejects. To the public, this doesn't always make sense. If there are discrepancies in policy directions between the offices, we don't do a good job explaining why one office regulated a chemical while another said it was OK."

In summary, the focus group participants said they wanted to know about any inherent weaknesses in the data which could affect their decisions. Having this type of information allows them to fully consider the impacts of their decisions and prepare to support those positions or decisions. Information risk managers seek on the weaknesses of the data seemed to encompass several different kinds of dimensions including these:

- Data gaps and significance of those gaps
- The degree to which an option will reduce risk rather than just shift it to another medium
- Stakeholders positions and likely reactions to the proposed options
- Facts about the efficacy and risks or unintended consequences of proposed alternatives or substitutes
- Conflicting decisions about the chemical made by other EPA offices and rationale for those decisions
- The proportion of risk attributable to a chemical that will actually be addressed in the proposed action

**2.6 Reactions to concrete pieces of risk information.** A focus group is an excellent vehicle for pretesting materials. A team of EPA risk assessors in the Durham and Research Triangle Park offices created seven risk information visuals designed to present uncertainties in a variety of formats. The focus group participants examined these examples. To duplicate the experience of an authentic briefing as much as possible, a videotape of several risk assessors in a briefing was created and shown to the group. The risk assessors explained the purpose of each figure. After each videotaped segment , the participants discussed their reactions to the piece. The videotape lasted approximately 10 minutes.

**Expected Number of People Experiencing Chest Discomfort One or More Times Per Year in Washington, D.C.**

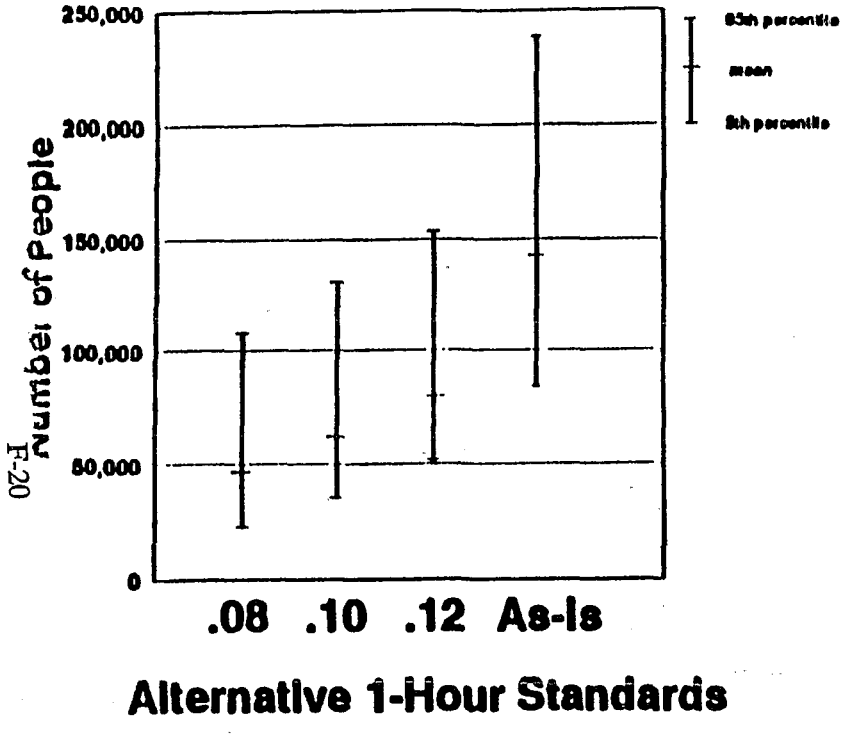



Figure 1

**Expected Number of People Experiencing Chest Discomfort One or More Times Per Year in Washington, D.C. Under Alternative 1-Hour Standards**

| Alternative 1-Hour Standards | Mean | 90% Credible Interval |
|---|---|---|
| As-Is | 145,000 | 82,000-240,000 |
| 0.12 ppm | 77,000 | 52,000-155,000 |
| 0.10 ppm | 60,000 | 35,000-130,000 |
| 0.08 ppm | 48,000 | 23,000-110,000 |

Figure 2

## Script for figure 1

"We've already discussed the two major inputs to the ozone risk model -- the exposure model and the probabilistic exposure/response relationships derived from the health effects data. We've also presented the uncertainties in each of these inputs to the model. The risk estimates shown in the figure result from combining these two inputs. The result is a set of estimates of the expected number of people experiencing chest discomfort one or more times per year under alternative air quality scenarios and standards. The scenarios are three different alternative one-hour ozone standards and an "as-is" situation representative of recent air quality in Washington D.C. The hash marks for each of these bars show the best mean or the best estimate of the expected number of people experiencing this effect. The mean, for example, is 145,000 under the "As-is" situation. The interval represents the 90% credible interval and can be interpreted as meaning there is only a one in 20 (or 5% chance) that the true number of people adversely affected is greater than 240,000 in the case of the "As-is" scenario, and only a 1 in 20 (or 5% chance) that the true number of people adversely affected is less than 82,000. The uncertainties indicated by the interval is due both to uncertainties represented in the exposure model and uncertainty in exposure-response relationships."

## Script for figure 2

"Another way to show this information is in a tabular format. Here we have the mean and the 90% credible intervals for the "As-is" situation and for each of the three standards. As you see, both the mean decreases as we go to more stringent standards, and the ranges decrease as well."

## Reactions to figures 1 and 2

The group members first viewed the videotaped briefings of figures 1 and 2, which showed two different ways of presenting estimates of the expected number of people experiencing chest discomfort one or more times a year under alternative air quality scenarios and standards. Figure 1 presented the information in a graph while the same information in figure 2 is presented in a tabular format.

The focus group participants liked both figures, and said that these are the types of visuals they would find useful. Although most group members preferred figure 2, there was consensus that they would like to have both figures included in a briefing packet. It is appropriate for risk assessors to present more than one visual conveying the same information in different formats, as the following comments illustrate:

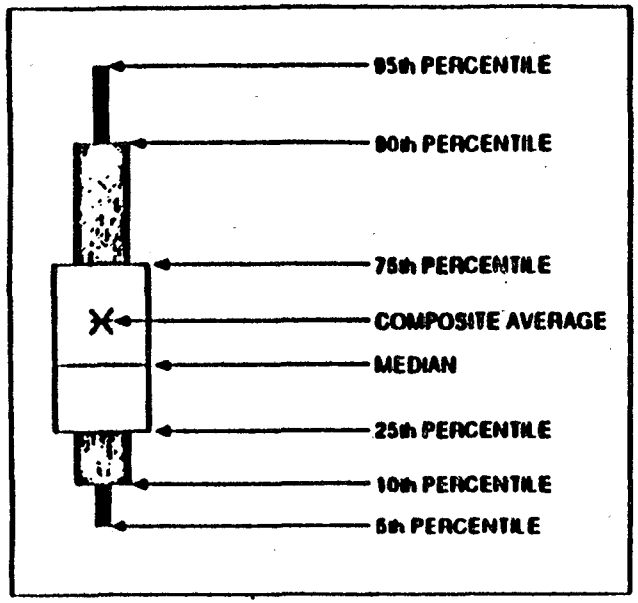
> "You need both of these figures. You can't tell this is [points to figure 1] 82,000 by looking at the graph, but you can by looking at the table."

> "There is a lot of information behind this. I would want both of these charts. figure gives you a sense of the magnitude, while figure 2 gives you specificity....Figure 1 gives a sense of proportion, while figure 2 provides more information."

Several participants said that the graph (figure 1) could be deceiving, since it suggests that going from the "as-is" situation to .12 is a big jump. The other conditions *seem* to go down by modest decrements, but that depends on whether "as-is" is .14 or .20. They indicated that "as-is" needs to have a number, stating:

> "Where is "as-is" relative to the other numbers? Is it .14 or is it .20? If it is .125 or .5, there's a big difference. I assume it is .14, but who knows?"

Participants noted that neither figure 1 nor figure 2 indicates how much confidence they can have in the numbers. Additionally, it is not clear if the numbers are based on modeling, monitoring, or expert opinion.


95th PERCENTILE
90th PERCENTILE
75th PERCENTILE
COMPOSITE AVERAGE
MEDIAN
25th PERCENTILE
10th PERCENTILE
5th PERCENTILE

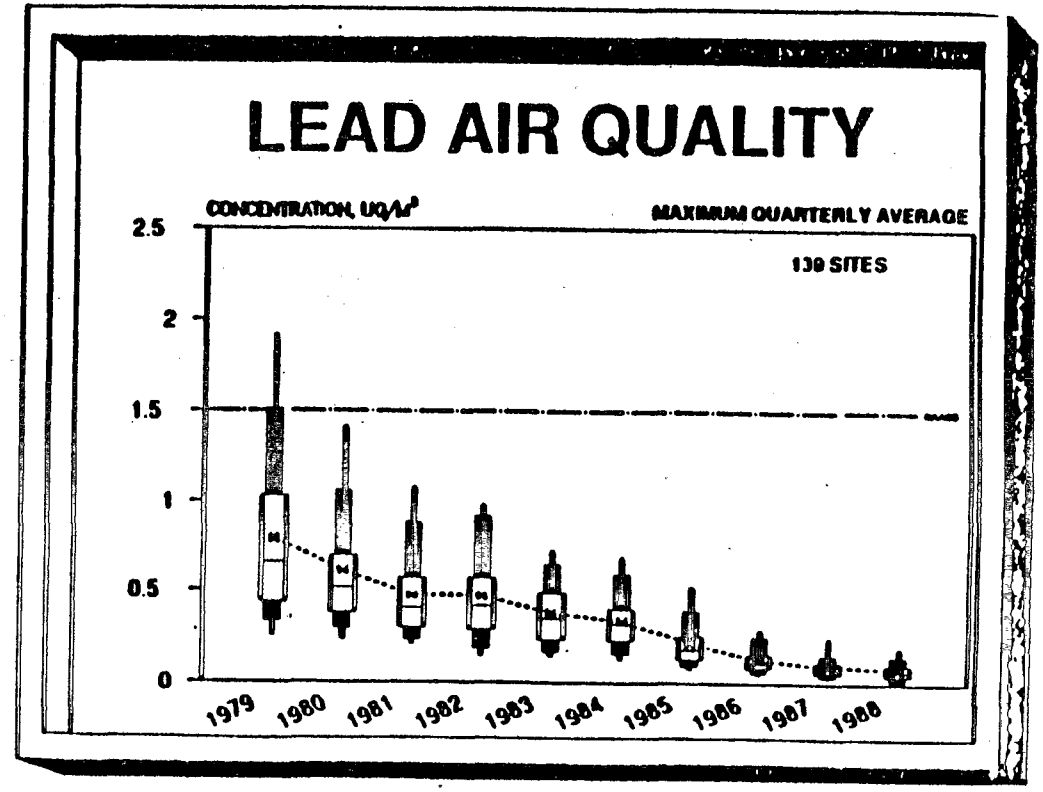

LEAD AIR QUALITY
CONCENTRATION, UG/M³
MAXIMUM QUARTERLY AVERAGE
139 SITES
2.5
2
1.5
1
0.5
0
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988

Figure 3

## Script for figure 3

"Ambient air quality data may be represented visually by a box plot. In the box plot what we have is the 95th percentile represented at the top of the plot. This indicates that 5% of the sites sampled were above this level. Then we have the 90th percentile in the first insert, then the 75th percentile, and the median which is the 50th percentile. Also shown is the composite average....Lead represents one of the true air quality successes in the United States. What we see from the box plots is that from the period 1980 to 1988 there was a broad wide-scale change in air quality standards with respect to lead. This is the ambient air quality standard. As you can see, in 1979 approximately 10% of the sites were above the ambient standard. That very quickly decreased with the phase-down of lead use and the variability decreased substantially also, so that in 1988 no sites exceeded the ambient air quality standard for lead."

## Reactions to figure 3

Figure 3 is a box plot which represents visually ambient air quality data. Most participants said they found box plots difficult to interpret. Furthermore, most thought that the box plot provided more information than they, as risk managers, need to develop a recommendation. One group member noted that lead (the subject of these figures) has a sufficiently rich data base to employ a box plot format. It is rare, however, to have enough information in the Air Office to be able to generate this type of chart. Most said they prefer a less detailed summary of the information. As two participants commented:

> "The basic message is that lead levels are going way down. We would get that from the dotted line. We don't need the rest of it."

> "The box plot gives you an overall impression -- a very striking one. It's effective. But if you were to try to make use of all this information of these eight different numbers for each year, it would be too much."

If risk assessors do provide box plots, they should do so in two stages, since the actual information is very hard to extract from this type of format. They should first show the dotted line by itself. Then they should show the box plot and orally summarize the variability information.

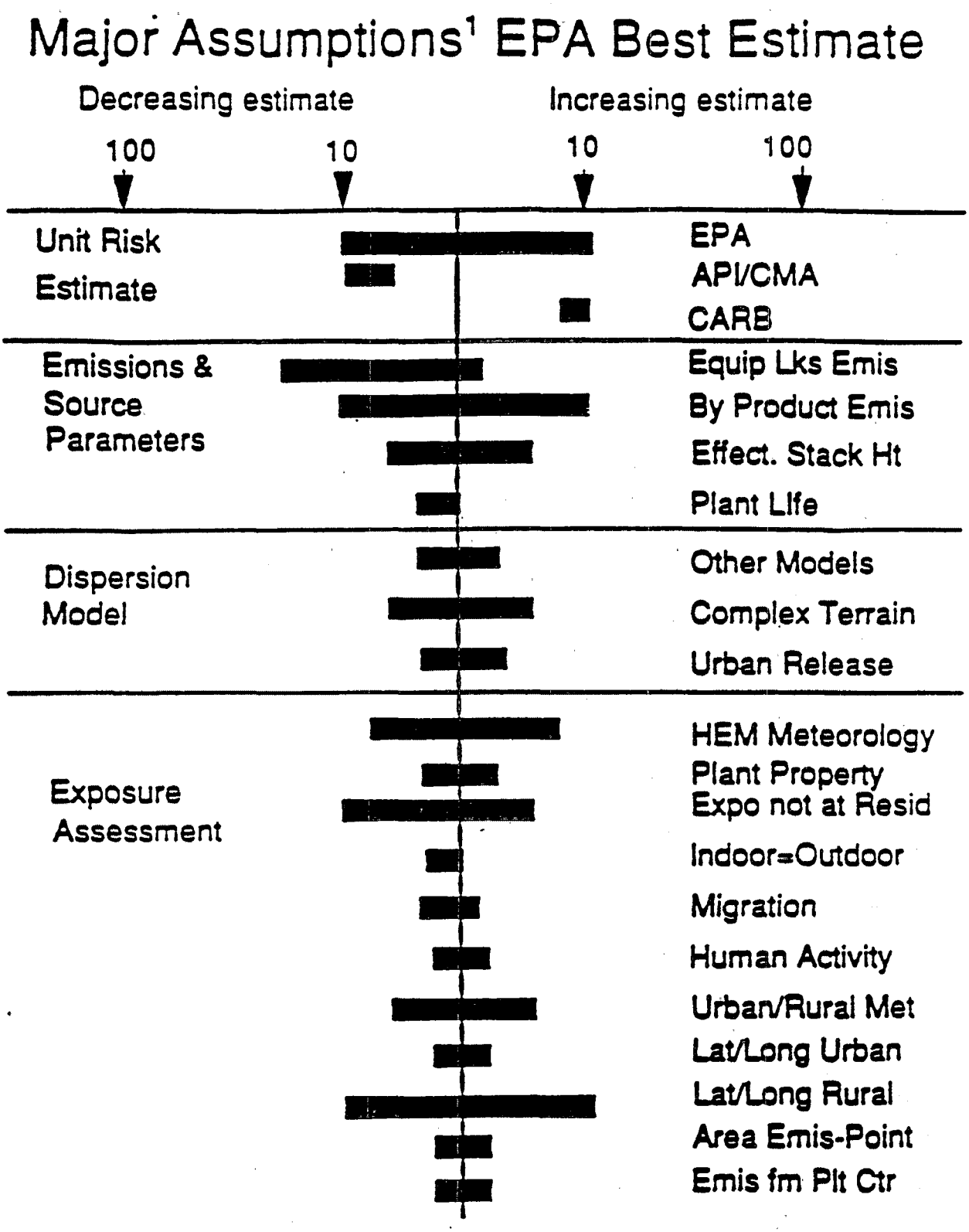

Major Assumptions[1] EPA Best Estimate
Decreasing estimate
Increasing estimate
100
10
10
100
Unit Risk Estimate
EPA
API/CMA
CARB
Emissions & Source Parameters
Equip Lks Emis
By Product Emis
Effect. Stack Ht
Plant Life
Dispersion Model
Other Models
Complex Terrain
Urban Release
Exposure Assessment
HEM Meteorology
Plant Property
Expo not at Resid
Indoor=Outdoor
Migration
Human Activity
Urban/Rural Met
Lat/Long Urban
Lat/Long Rural
Area Emis-Point
Emis fm Plt Ctr

Figure 4

Script for figure 4

"Now, let's look at some of the sources of uncertainty associated with the benzene estimates of the maximum individual lifetime risk. This chart represents some of the major sources of uncertainty split into 4 basic groups: the uncertainties surrounding the unit risk estimate, the uncertainty surrounding emissions and source parameters, dispersion model uncertainties, and uncertainties in exposure assessment.

The bars are assumed to be of uniform density in terms of uncertainty. We have shown them by orders of magnitude again with estimates increasing to 10 and 100 fold in this direction, and decreasing to 10 and 100 fold in this direction. EPA's estimate is represented by the vertical line here. If we look at the Unit Risk Estimate itself, we generally regard the reasonable uncertainty around that estimate to be an order of magnitude in either direction. Obviously there is some finite probability that the Unit Risk Estimate, or that the carcinogencity of the substance, in this case, Benzene, is zero.

But generally we feel our estimates are roughly accurate within an order of magnitude. On the other hand, the American Petroleum Institute and The Chemical Manufacturers Association have derived a unit risk factor for Benzene that is toward the lower bound of our reasonable estimate. This is based on a reanalysis of the same data set that was used to derive EPA's estimate. California Air Resources Board, on the other hand, using the mouse data set, which is only a portion of the animal data, has derived an estimate for the potency of Benzene that is much closer to the upper bound of EPA's estimate.

Looking at the emissions and source parameters, in this case the emissions from equipment leaks, in the original analysis we used estimates of emissions that were derived from emissions factors that EPA now believes to be very conservative. In this case, we are looking at potentially over-estimating the MIR for Benzene from anywhere up to a factor of about 20.

In the case of by-product plant emissions, we feel that there is a plus or minus condition here. That we are basically looking at about an order of magnitude in either direction potentially. That effective stack-height is somewhat symmetrically distributed around our own best-guess estimate. In the case of plant life, we feel that this represents an over-estimate of the risk since most plants tend to survive in-tact for about 20 to 50 years and our estimates are based on a lifetime risk of 70 years."

## Reactions to figure 4

Figure 4 is a chart showing some of the major sources of uncertainty associated with the benzene estimates of the maximum individual risk. Reactions to figure 4 varied. Some thought that it was useful to have the uncertainty bands around the major assumptions or parameters, while a few thought the chart conveyed too much information. They described it as "noisy for a high level decision maker" or "too busy." As one group member said:

> "I can't make very good use of this information. It's too much information for me. I'd like to see the first two sections in a chart and someone could tell me about the rest."

All participants said they would need someone to "walk them through this chart." Furthermore, they wanted a chart with this level of detail the day before the briefing so they would have ample time to study it.

There were several points of confusion in this figure. Almost everyone found the title, "Uncertainty Around MIR, Major Assumptions EPA Best Estimate" to be confusing. Most thought it was necessary to label the center line "EPA estimate." The terms at the top of the chart, "Decreasing estimate/Increasing estimate" were also somewhat confusing. Some group members suggested substituting "Lower value/Higher value," "Lower risk/Higher risk," "+/-," or "Less conservative/more conservative." In addition, participants wanted to know the source of this information.

They agreed that the chart could be revised as two separate charts: "We would like to see the bottom line in a simplified form, so we'd like a two-stage presentation." The first stage would present the EPA unit risk estimate, and the unit risk estimates of API/CMA and CARB. Such a figure could include a list of key assumptions that resulted in three such different risk estimates using the same data base. As one participant said:

> "My question is why are these estimates so different? Since two other groups came up with estimates very different than ours, they must have had very different assumptions. I want to know which assumptions were different. Then I would focus the discussion on why."

A second figure could include the EPA unit risk estimate, but eliminate the API/CMA and CARB estimates. Those who perceived this figure as "too busy" suggested eliminating the information concerning the dispersion model and exposure assessment. Others suggested leaving out parameters in which there was a high degree of certainty or consensus.

Group members were somewhat confused about which assumptions in this figure are most important. They wanted to know if there is a risk number with uncertainty bands that would reflect an overall estimate. They stated:

> "There are a lot of factors here but you don't know which ones are important. You can't tell from the chart which parameters really drive the overall estimate."

One risk manager, in particular, found this figure extremely useful because of the error bars surrounding each major area of uncertainty. To him, they indicate the range of scientific debate relative to the assumptions used in EPA's decision. He liked the figure's presentation of factors driving those estimates and EPA's position, relative to other groups:

> "This figure tells me what the key things were that were driving the uncertainties, what assumptions had to be made, and where we were relative to other groups. To me, it shows where we are vulnerable and where there is likely to be controversy."

## MANY PARAMETERS AFFECT RISK ASSESSMENTS

- SOME ARE VERY IMPORTANT
  - * POTENCY
  - * EMISSIONS
  - * STACK HEIGHT
  - * FENCELINE DISTANCE
- OTHERS ARE NOT
  - * VELOCITY
  - * TEMPERATURE

Figure 5

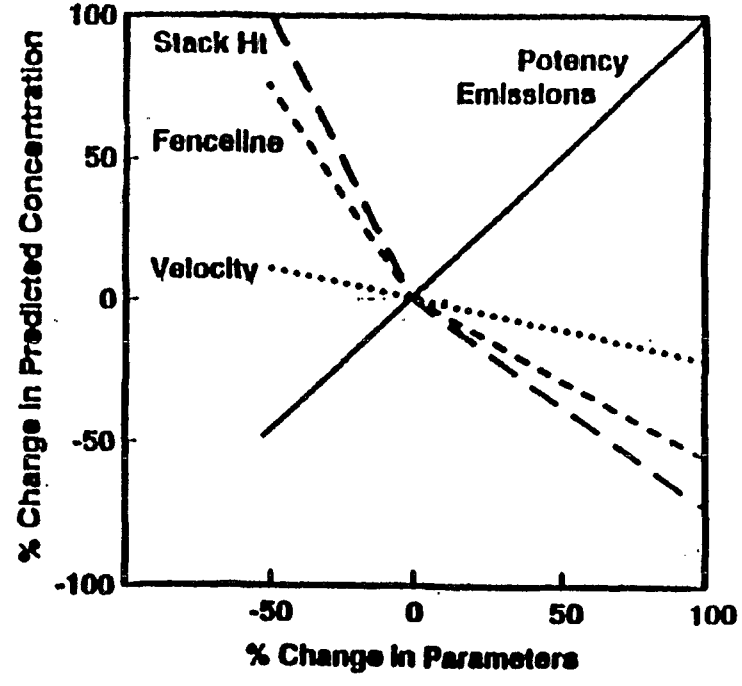



Figure 6

## Script for figure 5

"In estimating health risks from hazardous air pollutants we use an exposure and dispersion model which when coupled with a potency estimate provides us with the risk estimate. Many parameters are used in the risk estimation process. Some of these parameters are very important. These include the potency or in the case of a carcinogen, the unit risk estimate, the emissions estimate, the stack height or the height above ground level, and the distance to the fence line or the receptor.

Other parameters that are included particularly in a dispersion model are not as important. That includes the velocity of the emissions, and the temperature of the emissions.

## Script for figure 6

"We can look at the relative importance of different parameters visually or graphically in a plot such as this. Working from the base case of our standard default assumptions, what you see in the case of potency and emission (which are overlaid in this particular line), is that as we move from the default case to a doubling of those estimates, there is a direct relationship and a doubling of the predicted concentration. In the case of stack height and fence line distance, as you move from the default condition, in closer in terms of fence line distance, or to a shorter stack height, these are extremely important to the resulting risk estimates. In those cases also, as you move away from the default conditions, moving away in terms of fence line distance or moving to a taller stack height, these become somewhat less important. In the case of the velocity of the emissions, regardless of movement from the default condition to a less conservative or a more conservative assumption. It has very little impact on the resulting risk estimate. Looking at the entire picture, this tells us again that some parameters are critically important to the risk estimates while others are not."

## Reactions to figure 5

Figure 5 is a list of parameters which are important to the overall risk estimate and a list of factors which are not. The group agreed that figure 5, was too simplistic; it did not provide enough information for their decisions. They were confused about why "potency" and "emissions" would be shown on the same chart.

## Reactions to figure 6

Figure 6 presents graphically the relative importance of the different parameters which affect the risk estimate in a sensitivity analysis. The senior decision makers found Figure 6 too complex to understand easily. At first glance, they were overwhelmed by the level of detail. Decision makers at their level do not want or need that much detail. If the chart is used for technical analysis, the risk assessor should provide an introduction and then put up one line at a time to make the information manageable. They found having "potency" and "stack height" on the same chart was confusing. Some participants said they would prefer a bar chart to convey this information.

# Probability Density Function

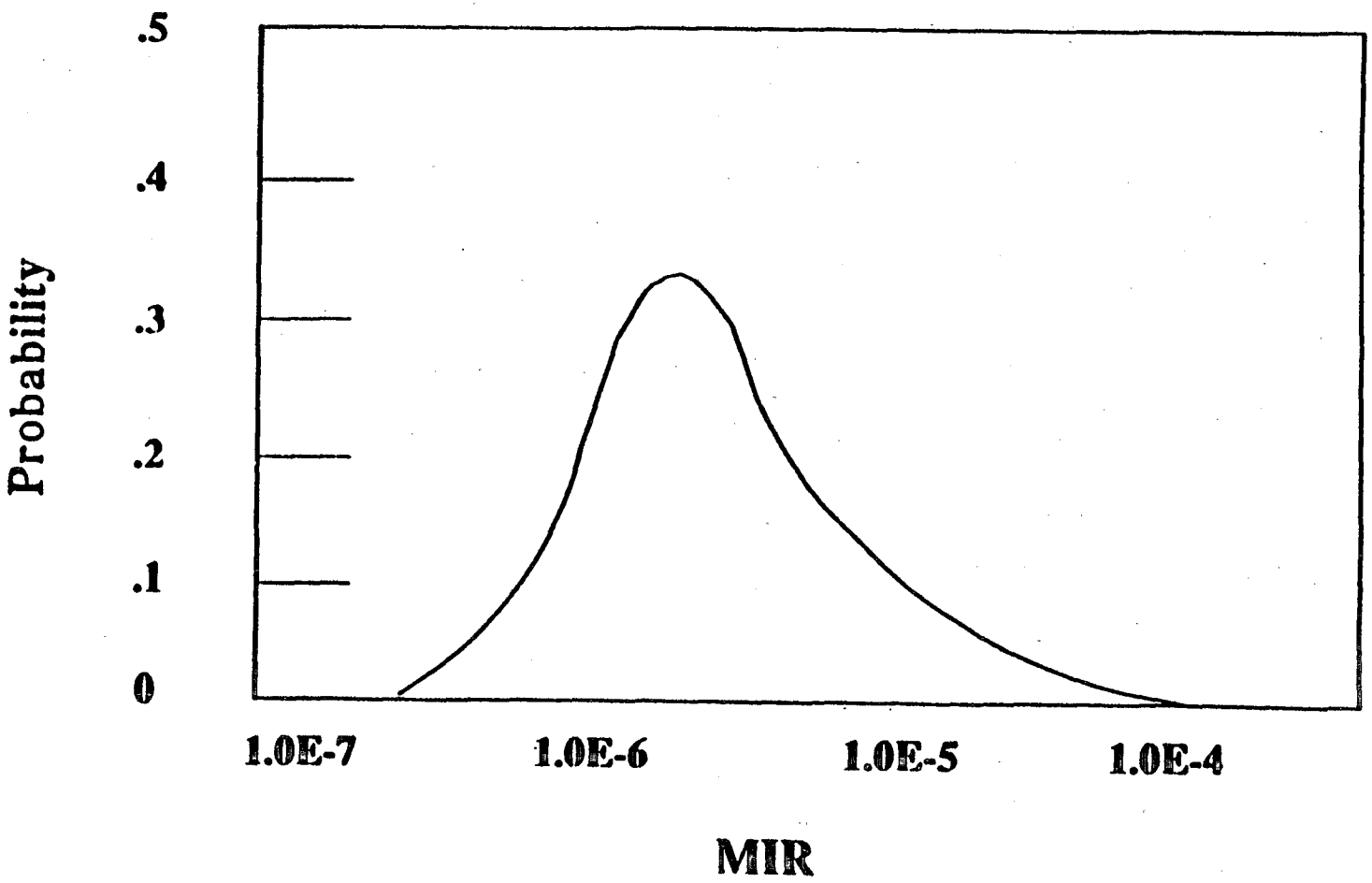



Figure 7

## Script for figure 7

"In the first hour we talked about how the Monte Carlo analysis is actually done. Now let's look at an output distribution. As we said earlier, we have allowed many of the critical parameters in the analysis to vary. In fact, we have represented them as distributions. This is a representative of the output distribution from the modeling run. There is a probability density function. As you can see, it's roughly a normal or log-normal distribution with probability on the vertical axis and the MIR represented on the horizontal axis. As you can see, the central tendency here of the distribution is somewhere in the neighborhood of 1 to 2 X 10-6."

## Reactions to figure 7

Figure 7 is an output distribution from a Monte Carlo analysis. The group liked the format showing a distribution. There was some confusion over whether this was a distribution of all the MIRs for the source category or just one MIR for a single source. Some said that there is not enough information presented. They would also need to know the underlying assumptions. Several said they would rather see cumulative probabilities.

## 3. CONCLUSIONS

In summary, OAQPS conducted a focus group of high-level EPA decision makers across different offices and programs throughout the Agency. Its purpose was to help risk assessors provide more useful risk information to policy makers. A number of key points emerged from the discussion concerning the kinds of information that decision makers would find most valuable. According to the eleven focus group participants, senior risk managers want a variety of qualitative information, in addition to quantitative risk measures, to help them when making regulatory decisions or recommendations. Based on the discussion of the group, the following key ideas could be useful to risk assessors in communicating risk information to EPA policy makers.

1. At the beginning of a briefing,

- Present an overview of why the action (e.g., regulation) under consideration is important.
- Identify who it is that cares about this issue.
- Present a picture of what the major stakeholders are saying about this issue.
- Discuss the positions of other EPA offices and other important constituents (e.g., OMB) on this issue.

2. When characterizing the risk of a chemical,

- Present information concerning the severity of the adverse health effect posed by the chemical. (Is it death or is it a runny nose?)
- Establish the extent to which scientists believe the chemical is really a hazard to humans.
- Tell what the level of confidence is in the data and in the numerical assessment of risk. (Tell who else has seen the information and who else agrees with it.)
- Explain where the data gaps are and tell how important those gaps are to the overall risk estimate.
- Highlight potential "high visibility data gaps" that are likely to become the focus of attention of groups outside EPA.

3. When discussing exposure,

- Define what population is at risk. (Is it the general population? Children? The elderly? Minority groups?)

- Estimate the number of people who are exposed to levels of concern and present the range of uncertainty around the exposure numbers.

4. When presenting risk management options,

- Present the legislative mandate
- Identify potential risk management options, including ones that have already been rejected along with reasons for rejection.
- Discuss each option in regard to its costs, ease of implementation, and likelihood of success for reducing risk.
- Clarify how much each option will reduce risk rather than merely shift it from one medium to another.
- Estimate what proportion of the risk from this chemical this action will actually address.
- Discuss the consequences of doing nothing.
- Review what has been done in previous similar situations.
- Mention studies in progress which could yield new and important information about this chemical.

5. In informing the decision maker about the weaknesses of data,

- Provide a sense of the uncertainties of the data.
- Discuss how effective any proposed alternative substitutes to this chemical would be in accomplishing their intended purpose.
- Discuss any risks, trade-offs, or unintended consequences of the proposed substitutes to this chemical.
- Bring to light decisions (especially conflicting decisions) that other EPA offices have made about this chemical.

In regard to graphics presented at briefings, the group had definite opinions about what types of charts and visuals are useful. In general, they did not want to see complex, busy charts with more detail than needed to make decisions. Box plots, for example, convey too much detail for risk managers' purposes. They liked the cumulative distribution format best.

The risk managers noted that since different people process information differently, it is appropriate for risk assessors to provide more than one format for presenting the same information in different ways. In general, they wanted visuals that are straightforward and clear. They did not want too much information in one chart. They recognized the need for accompanying written information and/or oral briefings. The following

comment illustrates the point that these decision makers do not expect a chart to convey all the important information at a glance:

> "At some point you have to be able to read. A chart summarizes what many paragraphs tell you. But if you end up putting everything in the chart you will have a paragraph....Charts usually don't give you a full sense of certainties and uncertainties of the science. I don't think there's a chart that's going to convey that. That's got to be done through written materials that accompany it or through an oral briefing. That's why you have briefings as opposed to just packaged charts. The charts are necessary but not sufficient."

When it is necessary to provide a more complex visual such as figure 4 on benzene, risk managers want to have the information before the briefing to study it, absorb the details, and think about the questions they want to ask. Especially for these complex charts, they recognize the need for a "good set of talking points and someone to walk you through it." They also suggested breaking up the information for complicated charts by showing it in stages.

At the end of the focus group, several participants endorsed holding a focus group of risk assessors to find out what information they believe senior EPA risk managers need to help them make well-informed decisions. Participants thought that results from such a group, in conjunction with information from their own focus group, would provide a balanced picture of how best to communicate risk to senior EPA risk managers.

# PRINCIPLES OF GOOD PRACTICE FOR THE USE OF MONTE CARLO TECHNIQUES IN HUMAN HEALTH AND ECOLOGICAL RISK ASSESSMENTS

**David E. Burmaster**
**Paul D. Anderson**

*Risk Analysis, Vol. 14, No. 4, 1994*

# Principles of Good Practice for the Use of Monte Carlo Techniques in Human Health and Ecological Risk Assessments

**David E. Burmaster[1] and Paul D. Anderson[2]**

*Received June 17, 1993; revised December 22, 1993*

We propose 14 principles of good practice to assist people in performing and reviewing probabilistic or Monte Carlo risk assessments, especially in the context of the federal and state statutes concerning chemicals in the environment. Monte Carlo risk assessments for hazardous waste sites that follow these principles will be easier to understand, will explicitly distinguish assumptions from data, and will consider and quantify effects that could otherwise lead to misinterpretation of the results. The proposed principles are neither mutually exclusive nor collectively exhaustive. We think and hope that these principles will evolve as new ideas arise and come into practice.

KEY WORDS: Probabilistic risk assessment; Monte Carlo.

## 1. INTRODUCTION

For over 50 years, Monte Carlo (MC) techniques have been used in physics, chemistry, and many other disciplines to compute difficult multi-dimensional integrals. One example of this use is to combine probability distributions for several input variables to estimate probability distributions for one or more output distributions.[12,14] The widespread use of Monte Carlo techniques in public health and environmental risk assessment promises significant improvements in the scientific rigor of these assessments. Because Monte Carlo methods are more computationally intensive than the "deterministic" or "point estimate" methods in common use today, some people have suggested that Monte Carlo analysis not be widely adopted at this time. We believe that this is an overreaction, but we recognize the need for safeguards and precautions to reduce mistakes and prevent abuses.

[1] Alceon Corporation, P.O. Box 2669, Cambridge, Massachusetts 02238-2669.

[2] Ogden Environmental and Energy Services, 239 Littleton Road, Suite 7C, Westford, Massachusetts 01886.

We propose 14 principles of good practice in this article to assist people in performing and reviewing probabilistic risk assessments, especially in the context of the federal and state statutes concerning chemicals in the environment. Monte Carlo risk assessments for hazardous waste sites that follow these principles will be easier to understand, will explicitly distinguish assumptions from data, and will consider effects that could otherwise lead to misinterpretation of the results. These proposed principles arise from years of experience conducting and reviewing MC risk assessments and from conversations with many knowledgeable people in manufacturing companies, consulting companies, law firms, universities, nonprofit organizations, and government agencies. We think and hope that these principles will evolve as new ideas arise and come into practice.

Before proposing the 14 principles, we agree that each risk assessment, whether deterministic or probabilistic in design, must have a clearly defined assessment end point[9] and must contain all the information such that a knowledgeable person can reproduce and then evaluate the analysis from the material presented in the final report.[13]



0272-4332/94/0800-0477$07.00/1 

## 2. THE PRINCIPLES

### 2.1. Principle 1

Show all the formulae used to estimate exposure point concentrations, exposure doses, toxic potencies, hazard indices, and/or incremental lifetime cancer risks. As for any risk assessment, show the formulae and the spreadsheets in the text, in tables, or in an appendix.

### 2.2. Principle 2

Calculate and present the point estimates of exposure and risk that are generated following the current deterministic risk assessment guidelines from the appropriate regulatory agency. The calculation of point estimates using standard techniques is a desirable first step in undertaking a MC risk assessment.

### 2.3. Principle 3

Present the results from univariate (or multivariate) sensitivity analyses of the deterministic calculations to identify the inputs suitable for probabilistic treatment and then discuss any variables not included in the sensitivity analysis. A typical risk assessment may require the specification of over 100 input variables. Only a few of these inputs drive the risk assessment in one or both of these senses: (i) The values of some inputs account for a dominant fraction of the predicted risks and/or (ii) the ranges of some inputs account for a dominant fraction of the range in the predicted risks. When using MC techniques, it is important to understand which inputs drive the predicted risk in both of these senses.

### 2.4. Principle 4

Restrict the use of probabilistic techniques to the pathways and compounds of regulatory importance to save time, money, and other scarce resources. For example, if a conservative, deterministic risk assessment shows that one pathway contributes $1 \cdot 10^{-8}$ incremental lifetime cancer risk, some two orders of magnitude below the typical threshold of regulatory concern of "one in one million" risk, then do not apply probabilistic methods to that pathway. This will save resources in the MC analysis without compromising its integrity or usefulness to a risk manager. Similarly, if some compounds contribute negligibly to the overall incremental lifetime cancer risk, then little need exists to undertake an expensive effort to estimate distributions for the Cancer Slope Factors (CSFs) or the Reference Doses (RfDs) for these compounds until such time as the US Environmental Protection Agency publishes distributions for CSFs and RfDs in their toxicological databases.

### 2.5. Principle 5

Provide detailed information on the input distributions selected. At a minimum, we suggest the following for each input distribution: (i) a graph showing the full distribution and the location of the point value used in the deterministic risk assessment and (ii) a table showing the mean, the standard deviation, the minimum (if one exists), the 5th percentile, the median, the 95th percentile, and the maximum (if one exists). In addition, the risk assessment should contain a 5- to 10-page justification of the selected distribution based on results in a refereed publication, from new developments, or from elicitation of expert judgment. For parametric distributions, discuss how the statistical process or the physical, chemical, or biological mechanism creating the random variable influences the choice of the distribution.[6]

### 2.6. Principle 6

Show, to the extent possible, how the input distributions (and their parameters) capture and represent both the *variability* and the *uncertainty* in the input variables.[1,4,8,9,13] [In this principle, we follow the growing usage of these terms in public health risk assessments: (i) variability ($V$) represents true heterogeneity in a well-characterized phenomenon which is usually irreducible through further measurement, while (ii) uncertainty ($U$) represents ignorance about a poorly characterized phenomenon which may be reducible through further measurements.] To the extent possible, it is important to specify the probability distributions for the input variables such that they capture both the $V$ and the $U$ inherent in each variable and permit $V$ and $U$ to be described and analyzed separately.[5,9,15]

### 2.7. Principle 7

Use measured data to inform the choice of input distributions whenever possible, after making sure that the data are relevant and representative to the population, place, and time in the study.[18] As appropriate for driving variables, undertake new field measurements to sup-

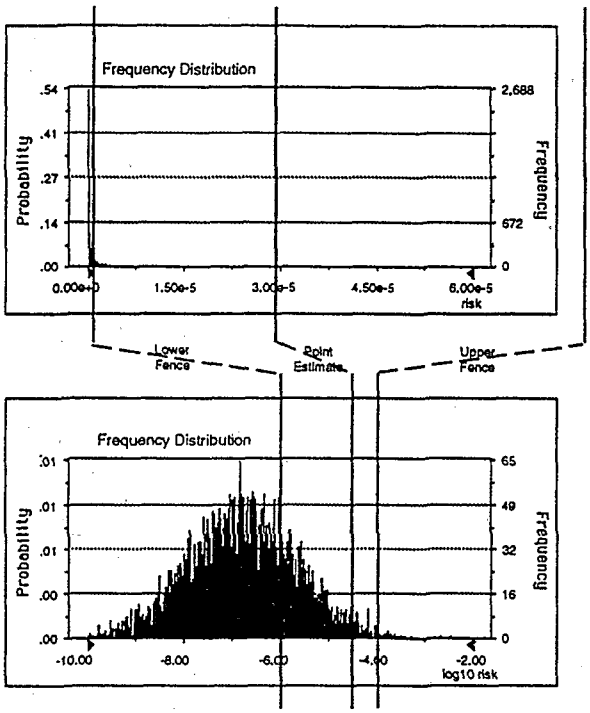



**Fig. 1.** Comparison of frequency distributions on linear and logarithmic scales.

ply missing information or to supplement partial information. If empirical measurements are not available for any reason, use and document accepted techniques—such as the Delphi method[3,13]—to estimate the input distributions for nonmeasured variables.

## 2.8. Principle 8

Discuss the methods and report the goodness-of-fit statistics for any parametric distributions for input variables that were fit quantitatively to measured data. Show plots of the parametric fits and the data on the same axes. Discuss the implications of any important differences. If any distribution was generated qualitatively or by expert judgment, discuss the techniques used.[18]

## 2.9. Principle 9

Discuss the presence or absence of moderate to strong correlations between or among the input variables. By strong correlation, we mean $|\rho| \geq 0.6$ or so. In many, but not all, practical situations, the absolute values of the correlations are less than 0.6. If so, the presence of moderate to strong correlations will have little effect on the central portions of output distributions[16] but may have larger effects on the tails of the output distributions. If it is possible that one or more moderate to strong correlations exist but no data are available from which to estimate them, perform Monte Carlo simulations with the correlations (i) set to zero and (ii) set to values considered high but plausible to learn if the possible correlations are important in the analysis. Display and discuss the results of these correlation sensitivity analyses and computational experiments, and state the practical effect, if any, of including or ignoring the correlations among the input variables.

## 2.10. Principle 10

Provide detailed information and graphs for each output distribution in the text and/or in an appendix. At a minimum, we suggest the following for each output variable: (i) a graph of the variable (in either log scale, linear scale, or both, depending upon the shape of the distribution) that clearly shows (a) the $10^{-4}$ risk and the $10^{-6}$ risk, or other allowable risk criteria, and (b) the point estimate of risk calculated by the deterministic method, and (ii) a table of the mean, the standard deviation, the minimum (if one exists), the 5th percentile, the median, the 95th percentile, and the maximum (if one exists). In Fig. 1, the histogram of estimated risk in the lower panel (on the log scale) gives a greater understanding of the variability in the output than does the histogram of the same results in the upper panel (on the linear scale). In Fig. 2, the histogram and the cumulative histogram in the upper and lower panels, respectively, display the variability of the output differently, but it is often useful to include both plots because each highlights a different aspect of the results. The graphs shown in Figs. 1 and 2 display the variabilities in the calculations, not the uncertainties.

## 2.11. Principle 11

Perform probabilistic sensitivity analyses for all of the key inputs represented by a distribution in the Monte Carlo analysis in such a way as to distinguish the effects of variability from the effects of uncertainty in the inputs. Display the results of these computational experiments in an appropriate graph.[9] The forms of the graphs will vary depending upon the method used to perform the probabilistic sensitivity analyses, but they should make clear which input variables contribute most

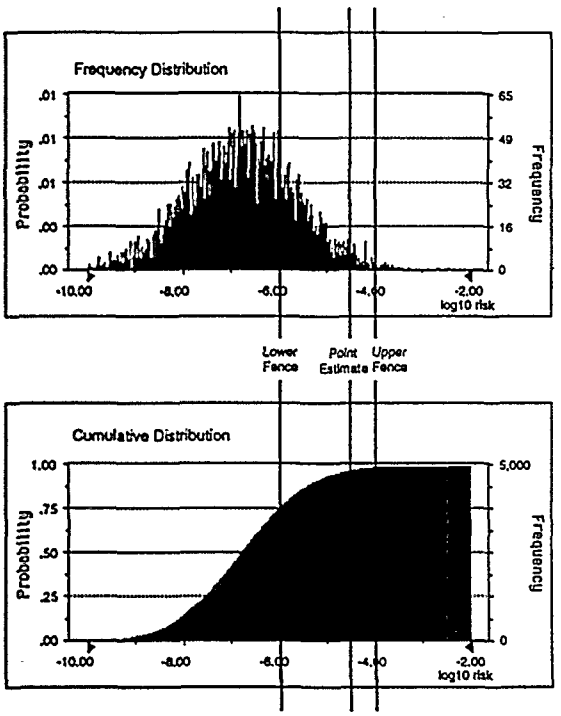



Fig. 2. Comparison of frequency distribution and cumulative distribution on a logarithmic scale.

strongly to the output variables. It is important to understand and display graphs showing which (groups of) input variables contribute most strongly to the (i) overall shape and location of the output distributions and (ii) the conservativeness, if any, created by point estimates in the deterministic analyses. For examples of these computational and visualization techniques, we recommend the papers by Ibrekk and Morgan,[10] Burmaster and von Stackelberg,[2] and Hoffman.[9]

### 2.12. Principle 12

Investigate the numerical stability of the (i) central moments (mean, standard deviation, skewness, and kurtosis) and (ii) the tails of the output distribution of the simulation. The tails of an output distribution are always less stable numerically than the central percentiles. In practice, the tails of the output distributions are more sensitive to changes in the tails of the input distributions. Because the upper tails of the output distributions often stabilize very slowly, the analyst should run enough iterations (commonly ≥10,000) to demonstrate the numerical stability of the tails of the outputs. If possible, the analyst should use software that includes Latin hypercube sampling (LHS) to help stabilize the tails of the outputs as quickly as possible. In addition, the analyst can and should discuss the sensitivity of the upper tails of the output distributions to changes in the upper tails of the input distributions. In practice, the changes in the tails of only a few input distributions contribute strongly to changes in the upper tail of the output distribution.

### 2.13. Principle 13

Present the name and the statistical quality of the random number generator used. Some well known commercial products have inadequate random number generators with short recurrence periods.[7] As the old computer saying goes, GIGO—"garbage in, garbage out." Too often, this inadvertently becomes "garbage in, gospel out." Call your software vendor and demand that she or he supply you with an audit from an independent testing laboratory that shows the strengths and limitations of generators and routines in the hardware and/or software. If you write your own specialty generator, include an appendix in your report listing the algorithm and the implementation, along with the results from a quality assurance audit.

### 2.14. Principle 14

Discuss the limitations of the methods and of the interpretation of the results. Be sure to acknowledge the source, the nature, and the possible effects of any unresolved sources of bias not explicitly included in the analysis, and indicate where additional research or measurements could improve the analysis.

## 3. DISCUSSION

Before an analyst undertakes a MC risk assessment, we hope that she or he will read widely in the growing literature on probabilistic risk assessment. We recommend reading and understanding the pathbreaking book *Uncertainty, a Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis*[13] as the minimum prerequisite. Morgan and Henrion—and many other authors—stress that the purpose and the objective of a study should guide its analysis. For example, at a hazardous waste site, there are important differences in objectives between a study to estimate baseline risks for

current conditions, a study to estimate risks for the reasonably foreseeable future conditions, and a study to estimate cleanup targets.

We have proposed these 14 principles of good practice as aids to performing or reviewing human health and ecological risk assessments done using MC techniques. While we favor the widespread use of MC techniques, we recognize the need for safeguards and precautions to reduce mistakes and prevent abuses. As proponents of the new methods, we hope that these proposed principles are general enough to show the standard of practice needed for conducting a MC assessment. We further hope that these ideas promote careful studies and innovation, which, in turn, create new insights and principles of good practice.

Several limitations apply to the ideas in this paper. First, the principles proposed are not mutually exclusive; some overlap with each other. Second, the principles proposed are not collectively exhaustive; for example, we have not proposed a principle concerning model uncertainty[13] nor one concerning the truncation of unbounded parametric input distributions (although the effects of truncation on percentiles and moments may be investigated through computational experiments and sensitivity analyses). Third, not all of these principles need apply to every study because not all of the principles are equally important in every situation. Fourth, the principles proposed are not inflexible recipes such as guidance manuals often present; we have instead tried to suggest the spirit of good practice without dictating a fixed and inviolate set of methods. Fifth, some of the principles are simply beyond the state of the art in some situations; for example, it is not now possible to fulfill all the proposed principles for a three-dimensional finite element model of time-varying ground water transport. Sixth, some of the principles are excessively burdensome for simple assessments. Notwithstanding all these limitations, we hope that the proposed principles will contribute to the quality of the MC studies undertaken. We further hope that these proposed principles will encourage others to refine these ideas to develop and publish new ones.

## ACKNOWLEDGMENTS

We thank Edmund A. C. Crouch, F. Owen Hoffman, Thomas E. McKone, Roy L. Smith, Alison C. Cullen, other colleagues, and two anonymous reviewers for helpful suggestions. Alceon Corporation and ENSR Consulting and Engineering supported this research.

## REFERENCES

1. K. T. Bogen. *Uncertainty in Environmental Risk Assessment* (Garland, New York, 1990).
2. D. E. Burmaster and K. von Stackelberg. "Using Monte Carlo Simulations in Public Health Risk Assessments: Estimating and Presenting Full Distributions of Risk," *J. Expos. Anal. Environ. Epidemiol.* 1(4), 491–512 (1991).
3. N. C. Dalkey. *The Delphi Method: An Experimental Study of Group Opinion*, RM-5888-PR (Rand Corporation, Santa Monica, CA, June 1969).
4. A. M. Finkel. *Confronting Uncertainty in Risk Management, a Guide for Decision-Makers* (Center for Risk Management, Resources for the Future, Washington, DC, Jan. 1990).
5. H. C. Frey. *Quantitative Analysis of Uncertainty and Variability in Environmental Policy Making* (AAAS/US EPA Environmental Science and Engineering Fellows Program, American Association for the Advancement of Science, Washington, DC, 1992).
6. D. B. Hattis and D. E. Burmaster. "Some Thoughts on Choosing Distributions for Practical Risk Analyses" (submitted for publication).
7. B. Hayes. "The Wheel of Fortune, The Science of Computing," *Am. Sci.* 81, 114–118 (1993).
8. F. O. Hoffman and J. S. Hammonds. *An Introductory Guide to Uncertainty Analysis in Environmental and Health Risk Assessment*, ESD Publication 3920 (Environmental Sciences Division, Oak Ridge National Laboratory, Oak Ridge, TN, Oct. 1992).
9. F. O. Hoffman. "Propagation of Uncertainty in Risk Assessments: The Need to Distinguish Between Uncertainty Due to Lack of Knowledge and Uncertainty Due to Variability," U.S. EPA/University of Virginia Workshop on When and How Can You Specify a Probability Distribution When You Don't Know Much, University of Virginia, Charlottesville, VA, 19–21 Apr. (1993).
10. H. Ibrekk and M. G. Morgan. "Graphical Communication of Uncertain Quantities to Nontechnical People," *Risk Anal.* 7, 519–529 (1983).
11. International Atomic Energy Agency. "Evaluating the Reliability of Predictions Using Environmental Transfer Models," Safety Practices Publications of the International Atomic Energy Agency, IAEA Safety Series, 100, 1–106 (1989).
12. B. J. T. Morgan. *Elements of Simulation* (Chapman and Hall, London, 1984).
13. M. G. Morgan and M. Henrion. *Uncertainty, a Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis* (Cambridge University Press, New York, 1990).
14. R. Y. Rubinstein. *Simulation and the Monte Carlo Method* (John Wiley and Sons, New York, 1981).
15. A. Shlyakhter and D. M. Kammen. "Sea-Level Rise or Fall," *Nature* 357, 25–7 (1992).
16. A. E. Smith, P. B. Ryan, and J. S. Evans. "The Effect of Neglecting Correlations When Propagating Uncertainty and Estimating Population Distribution of Risk," *Risk Anal.* 12, 467–474 (1992).
17. G. W. Suter II. *Ecological Risk Assessment* (Lewis, Chelsea, MI, 1993).
18. A. C. Taylor. "Using Objective and Subjective Information to Develop Distributions for Probabilistic Exposure Assessment," *J. Expos. Anal. Environ. Epidemiol.* 3(3), 285–298 (1993).

•U.S. GOVERNMENT PRINTING OFFICE:1996-549-001/60103