Benefits and Costs of Using Probabilistic Techniques in Human Health Risk Assessments --

with an Emphasis on Site-Specific Risk Assessments

David E. Burmaster, Ph.D.
Alceon Corporation
PO Box 382669, Harvard Square Station
Cambridge, MA 02238-2669
tel: 617-864-4300; fax: 617-864-9954
email: deb@Alceon.com

Introduction

Dictionaries define risk as the chance or probability of injury, damage, or loss (e.g., Webster's, 1970). Even though many other professions have long since adopted probabilistic frameworks using the Monte Carlo method pioneered during World War II, most human health risk assessments for exposures to carcinogenic compounds still use deterministic methods using point values for all variables. [See EndNote 1.] I believe that the benefits of adopting probabilistic methods for these studies outweigh the costs of the transition from current deterministic methods.

The Deterministic Framework

Building on the National Academy of Science's report titled *Risk Assessment in the Federal Government: Managing the Process* (NAS, 1983), the US Environmental Protection Agency (US EPA) has published many guidance manuals on deterministic risk assessment. In the deterministic framework, the incremental lifetime cancer risk, R, to a person exposed to a single carcinogenic chemical via a single pathway is a function of the exposure point concentration (EPC), represented by X_1; one or more exposure variables, represented by $X_2, \ldots, X_{N-1}$; and the appropriate cancer slope factor (CSF), represented by X_N. All of these variables represent real numbers, that is, point values.

$$R = f(X_1, X_2, \ldots, X_N) \qquad \text{Eqn 1}$$

This is a simple approach -- too simple really! It fails Albert Einstein's admonition: "Make things as simple as possible, but no more so." This formulation distorts the problem by simply ignoring the fundamental definition of risk as the probability of exposure and adverse outcome.

Moreover, US EPA's guidance manuals further distort this deterministic formulation of the problem (i) by selecting conservative or highly conservative values for many, if not all of the input variables, and then (ii) by compounding the conservatisms into an exaggerated point estimate through multiplication. A risk assessor who follows a guidance manual by combining several typical values, several conservative values, and several highly conservative values does not know where the estimated point value based on the compounded conservatisms falls on the overall distribution of risk. To the risk assessor's surprise, the estimated point value of risk may fall above the 95^{th}, the 98^{th}, or even the 99.9^{th} percentile of the overall distribution of risk (Cullen, 1994; Bogen, 1994; Burmaster & Harris, 1993). Here are two possible outcomes if the risk assessor does not know where a conservative point value of estimated risk falls on the overall distribution of risk:

- If a conservative point estimate of risk (i.e., one with compounded conservatisms) falls below some regulatory definition of maximum acceptable risk (now usually stated as a "bright line" test), then the risk assessor, risk manager, and members of the public can have confidence that the distribution of risk is truly acceptable. [See EndNote 2.]

- If a conservative point estimate of risk falls above the regulatory definition of maximum acceptable risk, then the risk assessor, risk manager, and members of the public do not know if the distribution of risk is truly unacceptable -- or if the apparently unacceptable risk is merely an artifact of the distortion inherent in the method and the guidance manuals.

<u>The Need for a New Framework</u>

Even within the deterministic framework, everyone should recognize and understand that the point value selected to represent a particular input variable is but one value from a range or from a distribution of possible values. For example, when selecting a single value to represent adult body weight, no one disputes the notion that different adults in a population have different body weights.

Within the deterministic framework, however, disputes do arise about how to select a policy-based point value to represent a particular input variable. Depending on the input variable, various people may recommend the arithmetic mean, the 95th upper confidence limit on the arithmetic mean, the median, the 90th percentile, or the 95th percentile of the range -- or some other point value chosen on the basis of policy, not science.

Many people think -- and I agree -- that the deterministic framework distorts risk assessment in irreparable ways, and that the selection of a single point value for each variable is misleading at best. Risk assessment needs a paradigm big enough and powerful enough to represent and analyze both variability and uncertainty inherent in the problem. Simply stated, variability deals with diversity in nature, while uncertainty deals with states of knowledge or ignorance (ideas adapted from Bogen, 1990).

- Variability represents knowledge of heterogeneity in a well characterized population. Variability is usually not reducible through further measurement or study. For example, a risk assessor knows that different adults drink different volumes of tap water each day, and the degree of variability would not be reduced by increasing the number of days of observation. [See EndNote 3.]

- Uncertainty represents ignorance about poorly characterized phenomena or models. Uncertainty is sometimes reducible through further measurement or study. For example, a risk assessor may know that a lognormal distribution describes the variability in the volumes of tap water ingested each day by a population of adults, but the risk assessor may be uncertain of the two parameters (e.g., geometric mean and geometric standard deviation) describing the lognormal distribution (except to within plausible ranges).

The deterministic framework does not address either of these fundamental issues in risk assessment, but the probabilistic framework can and does.

The Probabilistic Framework

Since risk is defined as the probability of an adverse outcome, probability, statistics, and the algebra of random variables provide the natural tools for quantifying and analyzing risk. In the probabilistic framework, the probability distribution of incremental lifetime cancer risk, $\underline{\underline{R}}$, to a person exposed to a single carcinogen via a single pathway is again a function of the exposure point concentration (EPC), the exposure variables, and the appropriate cancer slope factor (CSF), all of which are now probability distributions rather than real numbers:

$$\underline{\underline{R}} \quad = \quad f\,(\,\underline{\underline{X}}_1,\ \underline{\underline{X}}_2,\ \ldots\ ,\ \underline{\underline{X}}_N\,) \qquad\qquad \text{Eqn 2}$$

In Eqn 2, the function f remains the same as in Eqn 1, but now the N input variables and the single output variable are interpreted as random variables (see, e.g., Morgan & Henrion, 1990) denoted with the double underscores. [See EndNote 4.] Some subsets of the random variables $\underline{\underline{X}}_1$, $\underline{\underline{X}}_2$, ... , $\underline{\underline{X}}_N$ may have correlations and/or dependencies among them.

In mathematics, Eqn 2 is well defined. Its input distributions may describe the variability and/or the uncertainty inherent in the assessment at hand (Finkel, 1990). In practice, the probability distribution $\underline{\underline{R}}$ may be difficult or impossible to compute as a closed-form expression (Hoffman & Hammonds, 1993; Frey, 1992). However, with the advent of powerful computers, it is now quick, easy, and inexpensive to use Monte Carlo simulation or other methods to approximate the distribution $\underline{\underline{R}}$ to any degree of refinement (see, e.g., Rubinstein, 1981). [See EndNote 5.] Again, no one seriously disputes the mathematics or computations needed to approximate the probability distribution $\underline{\underline{R}}$ to a suitable tolerance.

Other disciplines have long included probability explicitly in risk assessments. For example, in safety engineering, the designers of chemical manufacturing plants, nuclear power plants, highways, and air traffic control systems -- to name but a few -- all include probability in the core of their analyses. Water resource engineers have pioneered the use of stochastic methods in water supply planning, i.e., rainfall and runoff, floods and droughts. Fisheries biologists use Monte Carlo simulation to study alternate management practices. Financial analysts on Wall Street and actuaries all use Monte Carlo simulation frequently.

At this point, let us acknowledge that probabilistic methods are well defined mathematically and have long been used in other professions and disciplines, including physics, chemistry, biology, engineering, economics, and finance. The simple fact that most human health risk assessments still use deterministic methods shows the need to persuade risk assessors, risk managers, regulatory agencies, and members of the public that the benefits of adopting probabilistic methods outweigh the costs of the transition.

The Benefits of the Probabilistic Framework

Probabilistic methods have key advantages over deterministic ones:

First, the probabilistic framework honors the definition of risk. Just as quantum mechanics places probability at its center, so must risk assessment.

Second, the probabilistic framework uses full information methods by including all the information available about the variability and the uncertainty inherent in the assessment. In the deterministic framework, the risk assessor discards most of the information about the variability and uncertainty in a phenomenon to pick one policy-based point value. [See EndNote 6.]

Third, the probabilistic framework reveals the compounded conservatisms inherent in the deterministic methods as currently used by regulatory agencies. In the probabilistic framework, risk assessors, risk managers, and members of the public all see the full range of variability and uncertainty, instead of being misled into thinking that exposure and risk are point values.

Fourth, probabilistic methods also reveal the nature and extent of the professional judgments in a risk assessment. As in the deterministic framework, professional judgment is the primary tool for quantifying uncertainty, but now the analyst must develop and reveal the ranges and distributions of the professional judgments in the analysis (see, e.g., Cooke, 1991).

Fifth, the probabilistic framework gives direct and indirect measures of the value of information. With the advent of inexpensive yet powerful desktop computers, we can now undertake probabilistic sensitivity analyses and other computational experiments to investigate which of many input distributions "drive" an analysis. We can now identify and isolate input variables with these characteristics: (i) the input has a high leverage on the final risk estimate and (ii) the input has a weak foundation in field measurements. Such an input is a strong candidate for additional field studies. Researchers in the "value of information" field use these techniques all the time, and we can readily adapt many of the techniques to human health risk assessment.

Sixth, probabilistic methods -- relying as they do on the full range of values that a variable may assume -- re-establish the now blurred boundary between risk assessment and risk management. Too often, a risk assessor working in the deterministic framework is required to use many high point values from a guidance manual -- point values chosen to exaggerate a problem so the risk manager can ignore the complexities and cost-effectiveness of a remediation.

Seventh, probabilistic methods ultimately save money. Yes, full information risk assessments may cost more than screening analyses using point values. But probabilistic assessments can derive less stringent -- yet fully protective -- cleanup targets, i.e., cleanup targets not based on compounded conservatisms. Since cleanup costs often rise asymptotically with decreasing cleanup targets, probabilistic assessments protect people and have high internal rates of return.

Last, and perhaps greatest of all, the probabilistic method's greatest advantage is that its output is a distribution of potential risk. And getting closer to the truth, no matter how difficult it may make our lives as risk assessors or risk managers, is preferable to the world of fiction created when distributions are replaced by single numbers. (Anderson, 1995).

<u>The Costs of the Probabilistic Framework</u>

Probabilistic methods entail certain costs associated with the maturation of a discipline:

First, probabilistic methods need more measured data than do deterministic ones. In the probabilistic framework, one or a few measurements are rarely enough to estimate the variability and the uncertainty in a random variable (see, e.g., Ott, 1995).

Second, probabilistic methods need many input distributions. For the probabilistic paradigm, the risk assessor needs to develop or tailor input distributions, not just use default distributions blindly. In a probabilistic risk assessment, the analyst selects and develops input distributions based on the variability and uncertainty in the situation, not on policy-based point values in a guidance document (see, e.g., Lipton et al, 1995).

Third, to work in the probabilistic framework, we risk assessors -- engineers, toxicologists, regulators and all -- need to learn new skills. Simply buying and running a computer program that does Monte Carlo simulation in a spreadsheet does not qualify any of us as a probabilistic risk assessor. To make the transition to the new paradigm, risk assessors need to take courses in probability, statistics, and simulation. First, it takes a change of world view to understand that the object of analysis is the full distribution of risk, not one or two points or summary statistics of the whole distribution. Second, it takes serious study to learn how to develop, manipulate, and interpret stochastic variables and equations.

Fourth, in the probabilistic framework, risk management decisions are harder. In the probabilistic paradigm, risk managers receive more information than in the deterministic paradigm. Now, she or he receives distributions for exposure and risk. While risk management in the deterministic paradigm usually consists of comparing a point value for estimated risk to a point value or range for acceptable risk, risk management in the probabilistic paradigm consists of judging a whole distribution of estimated risk as to its acceptability or unacceptability (see, e.g., Clemen, 1991). Given an estimated distribution for risk, the risk manager might use criteria or decision rules along these lines to render an opinion on the acceptability of the whole distribution of risk: (i) is the risk distribution highly skewed? (ii) is the median of the risk distribution less than 1 in a million?, (iii) is the arithmetic average of the risk distribution less than 1 in 100,000?, and (iv) is the 95^{th} percentile of the risk distribution less than 1 in 10,000? Thus the risk manager must consider the character, location, and spread of the whole distribution, not just one or two selected percentiles or summary statistics, when deciding whether a distribution of risk is acceptable for a population. By itself, any single point or summary statistic from a distribution of risk is misleading at best.

Fifth, in the probabilistic framework, risk communication hinges on the continued development of visual and graphical tools (see, e.g., Burmaster & von Stackelberg, 1991; Tufte, 1990; Ibrekk & Morgan, 1983; Tufte, 1983). While most people understand intuitively that all exposure and toxicity variables (e.g., body weight, the daily ingestion rate of drinking water, the number of days a person visits a park, and cancer slope factors) are random variables for different people in the population, few understand the mechanics of simulation or the algebra of random variables. Frankly, the algebra of random variables puts many people to sleep. As our profession moves to the probabilistic framework, we need to develop and use graphical and visual displays to convey the results. We risk assessors have to talk with risk managers and members of the public using visual images that all can understand.

Finally, as the risk assessment paradigm shifts from deterministic to probabilistic methods, we risk assessors must bear a burden as well: we must make sure that guidance manuals do not impede the growth and advancement of our discipline. In the probabilistic framework, we must not become prisoners of procrustean guidance manuals as we did in the deterministic framework.

Of course, many people see some or all of these "costs" as true benefits. As the switch to probabilistic methods requires more data, gives us full distributions of exposure, improves risk communications, and yields better, more cost-effective decisions, we risk assessors will look back in pride and wonder why the transition took so long to accomplish (Anderson & Graham, 1994).

Overall

In my view, the probabilistic paradigm honors the two basic tenets of risk assessment: First, the probabilistic paradigm builds on the fundamental definition of risk as the probability of an adverse outcome. Second, the probabilistic paradigm restores the distinction between risk assessment and risk management . Sadly, the deterministic paradigm violates both these tenets.

EndNotes

1. This essay discusses the estimation of risk to humans from exposures to carcinogenic chemicals. Some of the same arguments also apply to the estimation of risk to humans from exposures to noncarcinogenic chemicals or to the estimation of ecological risks to flora and fauna.

2. But they do not know the degree of (over) protection.

3. It is sometimes useful to study and then model the variability in a diverse population by disaggregating the population into a collection of more homogeneous subpopulations. For example, by disaggregating the population of people living in New York City into subgroups based on age, gender, and/or other factors, a researcher may learn important insights about the

variability in body weight in the overall population. However, the variability in body weight in the overall population does not decrease through disaggregation.

4. A random variable is a variable that can take any one of a range of values with a certain probability of occurrence. The range of values that the variable can take and the probability of those values are usually codified in a mathematical function called the probability function for that random variable. In practice, analysts can capture the range and probabilities in either of two (interchangeable) mathematical functions -- the first called the "cumulative distribution function" (CDF) and the second called the "probability density function" (PDF). The CDF is the integral of the PDF, so the two functions contain identical information.

5. Few people know how to multiply several random variables to create a new random variable (distribution) for the output variable. Even though multiplication of probability distributions is mathematically well defined, the calculation is tedious and involved. With the advent of powerful desktop computers, we now have commercial software packages that can perform the desired mathematical operations among random variables by a process called probabilistic simulation. The computer can estimate the output distribution by sampling the input distributions some 10,000 or 20,000 or more times and then assembling a list of the answers into an output distribution.

6. Is a deterministic risk assessment ever sufficient? Perhaps as a screening tool. Not every hazardous waste site or other risk assessment problem needs or deserves its own probabilistic assessment. For small sites with inexpensive remediation possibilities, it may make more sense to clean the property than to undertake a full probabilistic assessment. If a well-calibrated deterministic assessment indicates that risks are below the maximum acceptable point-value of risk, it makes little or no sense to go beyond the deterministic study. Similarly, if a deterministic assessment indicates that risks for certain exposure pathways or to certain exposed populations are below the maximum acceptable point-value of risk, it makes sense to exclude those pathways or populations from further study.

Acknowledgments

Kristen G. Edelmann, Kimberly M. Thompson, Paul Scott Price, Wendy H. Koch, Scott Ferson, James D. Wilson, F. Owen Hoffman, and Paul D. Anderson offered helpful criticisms on drafts of this essay.

Alceon Corporation funded the preparation of this essay.

References

Anderson, 1995

Anderson, P.D., 1995, personal communication

Anderson & Graham, 1994

Anderson, P.D. and J.D. Graham, 1994, Objections to Realism in Risk Assessment: Impediments to the Widespread Adoption of Distributional (Monte Carlo) Analysis, Annual Meeting of the Society for Risk Analysis, Baltimore, MD, December 1994

Bogen, 1994

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

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

Burmaster & Harris, 1993
Burmaster, D.E. and R.H. Harris, 1993, The Magnitude of Compounding Conservatisms in Superfund Risk Assessments, Risk Analysis, Volume 13, Number 2, pp 131 - 143

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

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

Cullen, 1994
Cullen, A.C., 1994, Measures of Compounding Conservatism in Probabilistic Risk Assessment, Risk Analysis, Volume 14, Number 4, pp 389 - 393

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

Hoffman & Hammonds, 1993
Hoffman, F.O. and J.S. Hammonds, 1993, Propagation of Uncertainty in Risk Assessments: The Need to Distinguish Between Uncertainty due to Lack of Knowledge and Uncertainty due to Variability, Risk Analysis, in press

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

Lipton et al, 1995
Lipton, J., W.D. Shaw, J. Holmes, and A. Patterson, 1995, Short Communication: Selecting Input Distributions for Use in Monte Carlo Simulations, Regulatory Toxicology and Pharmacology, Volume 21, pp 192 - 198

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

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

Rubinstein, 1981
Rubinstein, R.Y., 1981, Simulation and the Monte Carlo Method, John Wiley & Sons, New York, NY

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

Webster's, 1970
Webster's New World Dictionary of the American Language, D.B. Guralnik, ed, World Publishing Company, New York, NY