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-

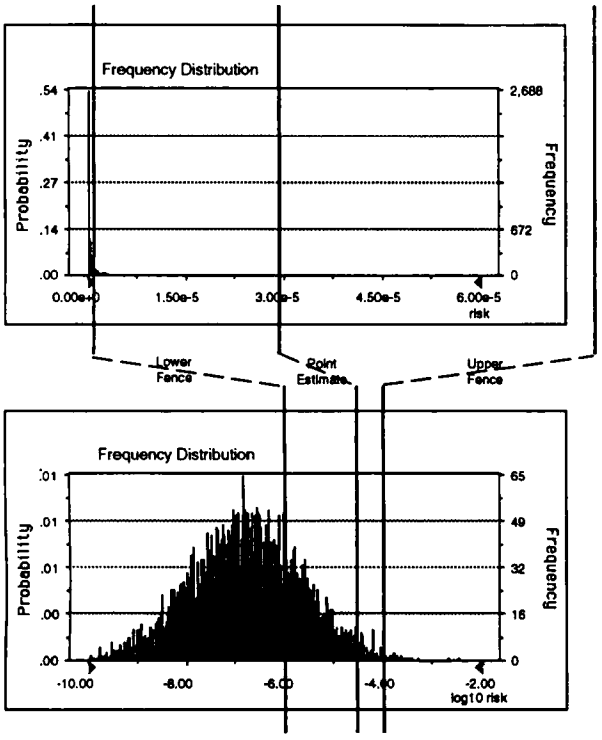

Fig. 1. Comparison of frequency distributions on linear and logarithmic scales.

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

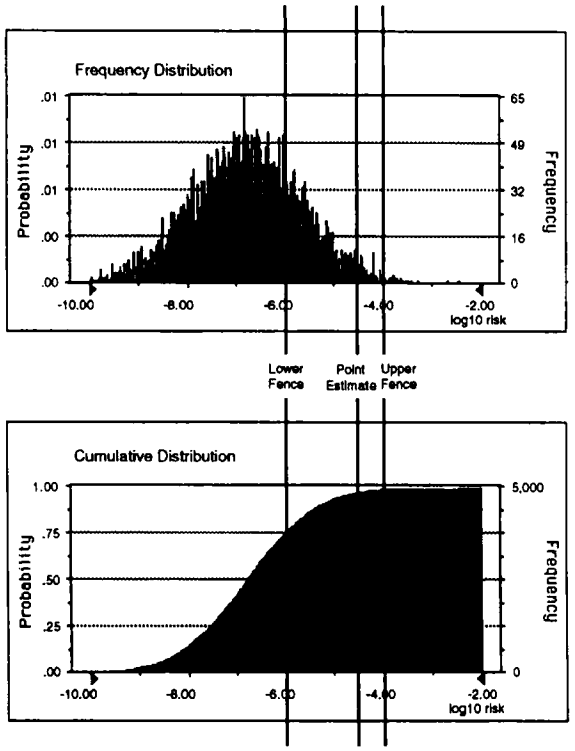

Fig. 2. Comparison of frequency distribution and cumulative distribution on a logarithmic scale.

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).