

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/15398068>

Propagation of uncertainty in risk assessments: The need to distinguish between uncertainty due to lack of knowledge and...

Article in Risk Analysis · November 1994

Source: PubMed

CITATIONS

279

READS

2,210

2 authors, including:



[Owen Hoffman](#)

Oak Ridge Center For Risk Analysis

151 PUBLICATIONS 2,359 CITATIONS

SEE PROFILE

Propagation of Uncertainty in Risk Assessments: The Need to Distinguish Between Uncertainty Due to Lack of Knowledge and Uncertainty Due to Variability

F. Owen Hoffman¹ and Jana S. Hammonds¹

Received December 30, 1993

In quantitative uncertainty analysis, it is essential to define rigorously the endpoint or target of the assessment. Two distinctly different approaches using Monte Carlo methods are discussed: (1) the end point is a fixed but unknown value (e.g., the maximally exposed individual, the average individual, or a specific individual) or (2) the end point is an unknown distribution of values (e.g., the variability of exposures among unspecified individuals in the population). In the first case, values are sampled at random from distributions representing various "degrees of belief" about the unknown "fixed" values of the parameters to produce a distribution of model results. The distribution of model results represents a subjective confidence statement about the true but unknown assessment end point. The important input parameters are those that contribute most to the spread in the distribution of the model results. In the second case, Monte Carlo calculations are performed in two dimensions producing numerous alternative representations of the true but unknown distribution. These alternative distributions permit subject confidence statements to be made from two perspectives: (1) for the individual exposure occurring at a specified fractile of the distribution or (2) for the fractile of the distribution associated with a specified level of individual exposure. The relative importance of input parameters will depend on the fractile or exposure level of interest. The quantification of uncertainty for the simulation of a true but unknown distribution of values represents the state-of-the-art in assessment modeling.

KEY WORDS: Uncertainty; variability; lack of knowledge; Monte Carlo methods.

1. INTRODUCTION

Present-day risk assessments for Superfund sites are generally conducted by using point estimates for parameter values and applying these values in an additive series of multiplicative equations.⁽⁸⁾ These calculations are designed to produce a reasonable approximation of exposure and health risk (probability of a harmful effect) to a maximally exposed individual. Although the severity of the risk is defined numerically, few risk assessments performed in support of Superfund regulations

contain a formal analysis of uncertainty. An uncertainty analysis, if performed at all, is usually restricted to a qualitative statement of confidence in the result; for instance, uncertainty in the point estimate that is less than one order of magnitude (a factor of 10) is considered "low," uncertainty in the point estimate greater than one order of magnitude but less than two orders of magnitude (a factor of 100) is considered "moderate," and uncertainty that exceeds two orders of magnitude is considered "high."⁽⁸⁾ Unfortunately, these qualitative statements of uncertainty are difficult to assess, let alone defend, particularly when the assessment involves potential exposure to several contaminants transferred over a number of different pathways.

¹ SENES Oak Ridge, Inc., Center for Risk Analysis, 677 Emory Valley Road, Suite C, Oak Ridge, Tennessee 37830.

A more defensible approach is to perform a quantitative analysis of uncertainty using either analytical or numerical techniques to propagate uncertainty in the components of the risk assessment equations into an assessment of uncertainty in the overall result. However, the manner in which these techniques are applied and interpreted will vary depending on the definition of the assessment endpoint (i.e., the target of the assessment question).

For any given risk assessment, the definition of the assessment endpoint is critical. This definition determines the relevancy of available data, the individuals or populations at risk, and whether or not it is necessary to distinguish between probabilistic statements about the variability of risks among individuals and uncertainty due to lack of knowledge about fixed but unknown quantities.

2. The Distinction Between Variability and Uncertainty: Type A Versus Type B Uncertainty

If the assessment endpoint is a fixed quantity (such as the risk to a specific individual, a maximally exposed individual, or an average individual of a specific population subgroup), the uncertainty analysis provides a statement of confidence that the true quantity will be within certain limits. The probability distribution obtained from uncertainty analyses using Monte Carlo simulation represents a range of "degrees of belief" that the true but unknown value is equal to or less than any value selected from the distribution (Fig. 1). This statement of confidence accounts for multiple sources of uncertainty, including uncertainty associated with the model structure and the presence, variability, and representativeness of data. Uncertainty about a quantity that is fixed (or deterministic) with respect to the assessment endpoint is called Type B uncertainty in Safety Series No. 100 of the IAEA.⁽⁶⁾ When the assessment endpoint is a fixed quantity, distributions of values obtained from repeated observations represent uncertainty of Type B because the "true" value is still an unknown quantity. The observations can be used to construct a confidence interval for which there is a given percent chance of bounding the true value.

Recently, EPA has been requesting that risk assessments target the upper 95th percentile of a population of potentially exposed individuals.⁽⁹⁾ This request requires simulation of the distribution of actual individual exposures or risks in the population. When the assessment endpoint is a distribution of actual exposures or risks (but the exposure to specific individuals in the

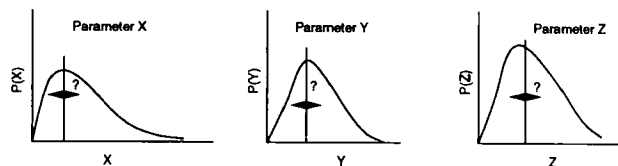
ASSESSMENT WITH AN ENDPOINT THAT IS A FIXED BUT UNKNOWN QUANTITY
(i.e. those which only have a single true value)

UNCERTAINTY IS ENTIRELY OF TYPE 'B'

EXAMPLE:

Risk to a specific individual from a given release, or
Risk to the Average Member of a Critical Group

Parameter X: Source Term parameter that is fixed but unknown
Parameter Y: Environmental Transport parameter that is fixed but unknown
Parameter Z: Dosimetric and risk conversion parameter that is fixed but unknown



THE TRUE VALUE OF EACH PARAMETER IS UNKNOWN;
EACH DISTRIBUTION REPRESENTS A SET OF POSSIBLE ALTERNATIVE
VALUES FOR THAT PARAMETER

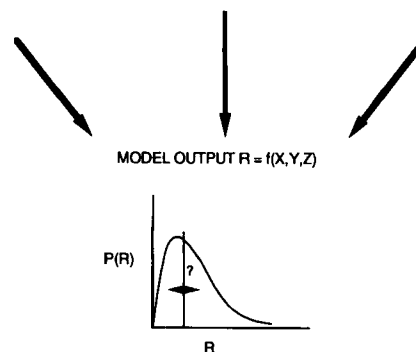


Fig. 1. Use of a Monte Carlo approach to estimate Type B uncertainty when the assessment endpoint is a fixed but unknown quantity.

population remains unknown), the uncertainty is of Type A.⁽⁶⁾ However, the true mean, variance, and shape of this distribution are "fixed" with respect to the assessment endpoint. If these quantities are unknown, uncertainty of Type B is also present.

When the assessment endpoint is a distribution of actual exposures or risks, it is necessary to make distinctions between Type A and Type B uncertainties. Such a distinction permits a confidence interval to be estimated for the true 95th percentile of a potentially exposed population (or for any other fractile of the true but unknown distribution of exposures or risks).

To distinguish between Type A and Type B uncertainty, a Monte Carlo simulation must be applied in two dimensions. First, numerous sets of alternative values are obtained from marginal probability density functions (PDF's) representing subjective degrees of belief about quantities that are fixed but unknown with respect to the assessment endpoint. Fixed quantities include parameters that do not vary with the assessment endpoint, such as the total amount of the contaminant released. Fixed quantities also include the mean, variance, and shape of

ASSESSMENT WITH AN ENDPOINT THAT IS A DISTRIBUTION OF ACTUAL INDIVIDUAL EXPOSURES OR RISKS

EXAMPLE: Distribution of risks to unspecified individuals in a population

Parameter X: Source Term parameter that is fixed but unknown

Parameter Y: Environmental Transport parameter that varies per exposed individual

Parameter Z: Dosimetric and risk conversion parameter that varies per exposed individual

VARIABLES ARE EITHER TYPE 'B' OR A COMBINATION OF TYPE 'A' AND TYPE 'B' UNCERTAINTY

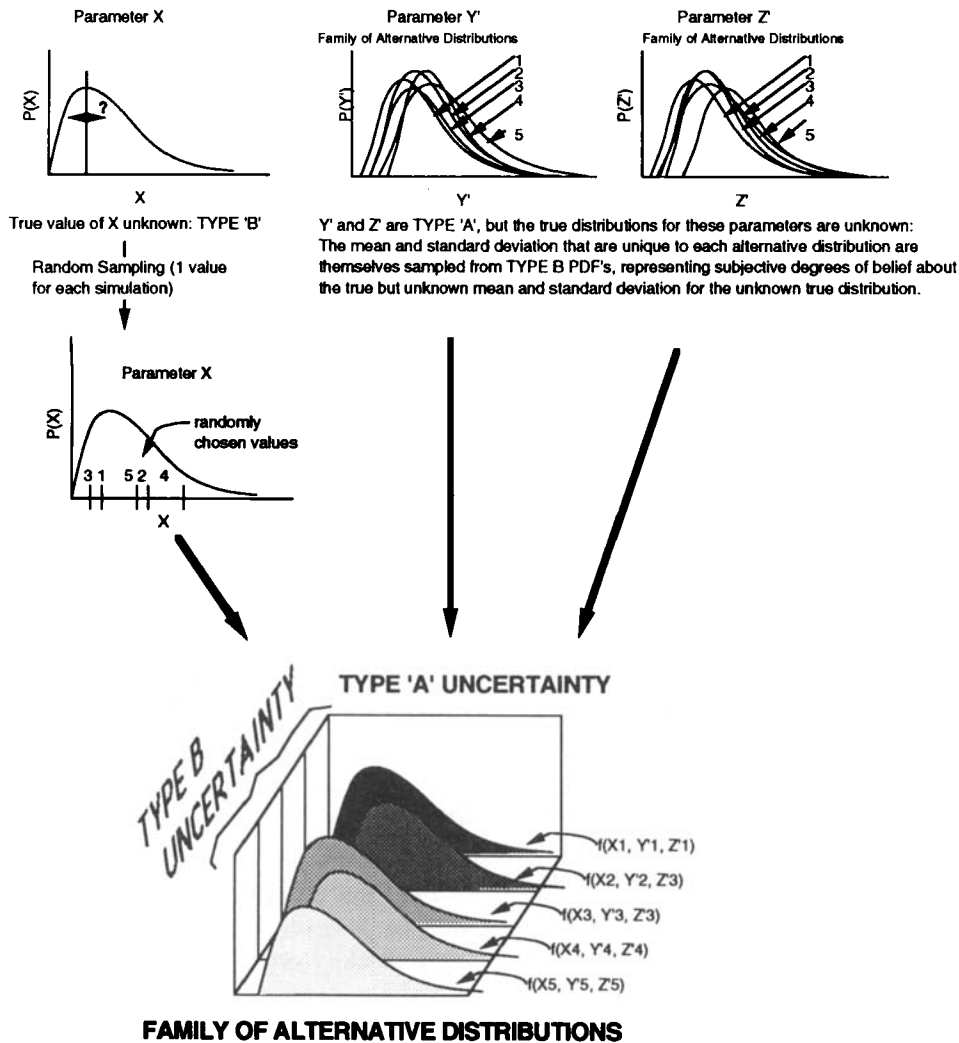


Fig. 2. Use of a Monte Carlo approach to distinguish between Type A and Type B uncertainty when the assessment end point is a true but unknown distribution of values representing variability among unspecified individuals in an exposed population.

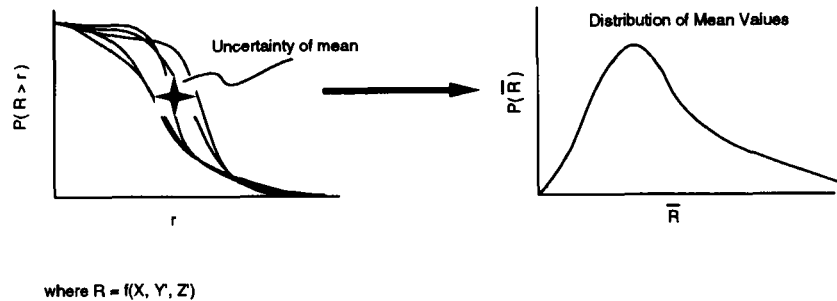
those parameter distributions that describe variability among individuals, as well as values that describe correlations among these parameters. The alternative sets of fixed values represent Type B uncertainty.

Second, for each alternative set of fixed values, Monte Carlo procedures are used to simulate alternative distributions of parameter values (marginal PDF's) that determine alternative expressions of the variability of in-

dividual risks (joint PDF's), each with its own unique mean, variance, and shape (Fig. 2). Each of these joint PDF's is thus an alternative representation of Type A uncertainty. The entire set of alternative distributions represents Type B uncertainty due to the fact that the true distribution is unknown. The alternative distributions are used to obtain subjective confidence intervals for the unknown risk at any given fractile. Subjective

**TWO INTERPRETATIONS FOR ALTERNATIVE REALIZATIONS OF ASSESSMENT
ENDPOINTS REPRESENTING TYPE 'A' UNCERTAINTY**

- ① TYPE 'B' uncertainty if the estimate of the true mean does is of interest



- ② TYPE 'B' uncertainty if the estimate of the true distribution is of interest

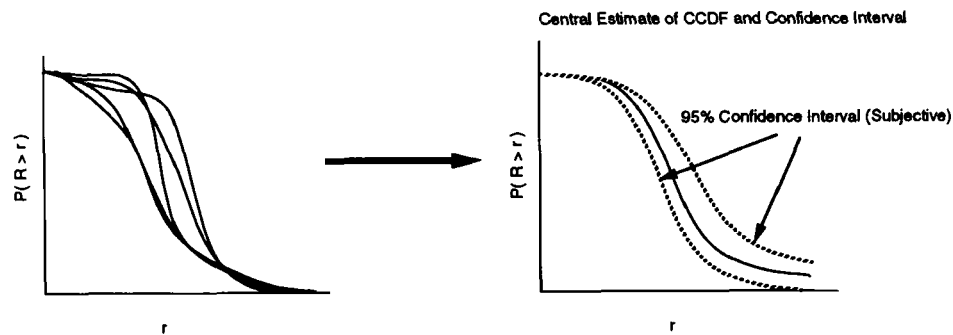


Fig. 3. Numerous alternative distributions produced through Monte Carlo simulation of Type A and Type B uncertainty can be used to derive confidence intervals for the mean value and any fractile of the true but unknown distribution.

confidence intervals can also be obtained for the unknown fractile at any given value of risk (Fig. 3). The order of importance for the parameters that contribute most to the confidence interval for a fractile will depend on the fractile of interest.⁽⁶⁾ Additional readings on this issue can be obtained from a number of authors.^(1,3-5,7)

Example

It should be noted that the example presented in this paper is entirely hypothetical and is for demonstration purposes only.

Let us assume that there has been an accidental spill of mercury in a small lake. The problem is to determine the Hazard Quotient (HQ) for the upper 95th percentile of the population of potentially exposed individuals from the ingestion of contaminated fish. Table I provides information required to solve this problem.

Solution

For this problem, we use a simplified form of an equation presented in the EPA Superfund guidelines for human health risk assessment:⁽⁸⁾

$$HQ = C_f \times I \times (BM)^{-1} (RfD)^{-1}$$

where

HQ = the Hazard Quotient (unitless)

C_f = the concentration of the contaminant in fish (mg/kg)

I = the ingestion rate of fish after consideration of the exposure duration, exposure frequency, and averaging time (kg/day)

BM = the body mass of the human (kg)

RfD = the chemical-specific reference dose (mg/kg-day)

Table I. Information About the Variability (Type A) and Uncertainty (Type B) of the Model Parameters Required to Solve the Example Problem

Parameter		C_t (mg/kg)		I (kg/day)		BM (kg)		
Type A	Dist.	log-Normal		log-Normal		Truncated normal ^a		RfD
		Mean	STDEV	Mean	STDEV	Mean	STDEV	(mg/kg-day)
Type B	Dist.	Triangle	Triangle	Triangle	Triangle	Triangle	Triangle	log-Tri.
	Min.	5.00E-2	2.50E-2	2.00E-2	2.20E-03	60	7	9.50E-5
	Mode	7.10E-2	3.43E-2	6.50E-2	2.80E-03	70	10	3.00E-4
	Max.	1.00E-1	5.00E-2	1.30E-1	4.40E-03	90	14	3.00E-3

^a Truncated at zero.

When the assessment end point is the 95th percentile of the exposed population, the variability of the mean concentration of the contaminant in fish and of the body mass and the daily ingestion rate for each exposed individual must be considered (Type A). Type B uncertainty also exists because the true values of the means and standard deviations of these three assessment parameters are unknown. In this example, the RfD for mercury is assumed not to vary from individual to individual; therefore, only Type B uncertainty is considered in this parameter.

The first step is to express uncertainty associated with the unknown values for the deterministic or non-varying quantities (Type B uncertainty). A log-triangular distribution was assigned to represent the uncertainty (due to lack of knowledge) about the true but unknown value of the RfD (Table I). Triangular distributions were assigned to the unknown means and standard deviations used to describe the stochastically varying parameters C_f , BM, and I . From these distributions that represent Type B uncertainty, 59 values of the reference dose and 59 values of the means and standard deviations of C_f , BM, and I were obtained using simple random sampling (SRS) in a Monte Carlo simulation.

The stochastic variability in C_f , I , and BM are determined by the values of the means and standard deviations that define the Type A distributions that are relevant to variability in the assessment end point (Table I). From each of the Type A distributions for C_f , I , and BM, one value is obtained using Latin hypercube sampling (LHS). These values are then combined with one of the 59 values selected for the RfD to simulate one value for the HQ for mercury ingestion. This process is repeated 100 times, without changing the value for the RfD or the values for the mean and standard deviation describing Type A uncertainty in each of the three parameters. Thus, a single Type A distribution (a joint PDF_v) is produced which is composed of 100 values

representing the HQ for individuals randomly selected from an exposed population.

Next, a new value is obtained from the 59 random values selected for the RfD; new values are also obtained for the mean and standard deviation of C_f , BM, and I ; and another distinct set of 100 results is produced representing an alternative set of values of HQ for an alternative set of individuals randomly selected from an exposed population. This process is repeated until there are 59 unique Type A distributions, each composed of 100 values representing alternative expressions of the variability of HQ among individuals in an exposed population (Fig. 4). With 59 alternative representations of a true but unknown distribution of HQ, it is possible to obtain a subjective confidence interval at any fractile of interest. For this example, the 90% subjective confidence interval for the HQ occurring at the 95th percentile of the population is 0.14 to 1.10 (Fig. 4). [End of example.]

In the example below, the 90% subjective confidence interval was conservatively estimated using statistical tolerance limits.^{2,6} Based on the minimum and maximum values obtained from the 59 simulations of

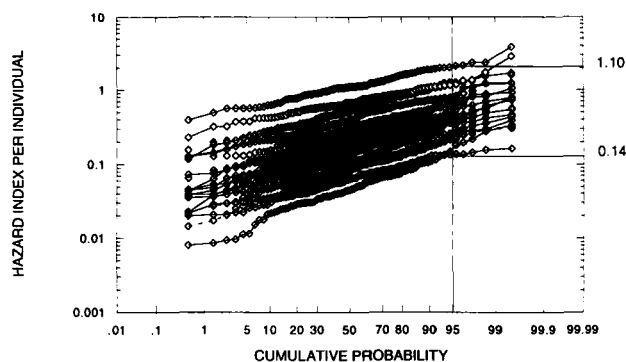


Fig. 4. Presentation of results from an uncertainty analysis containing Type A and Type B uncertainty.

the 95th percentile of the HQ. This relatively small sample size was used for computational convenience. This method provides a distribution-free statistical tolerance limit in which there is 95% confidence that the minimum and maximum values from an SRS of 59 are not underestimates of the 90% subjective confidence interval that would have been obtained using a much larger sample size (many thousands). This statistical tolerance limit is independent of the number of uncertain parameters in the model.⁽⁶⁾ An alternative approach would be to use a much larger sample size to produce a very large number of alternative realizations of the Type A distribution for HQ. In this case, the 90% subjective confidence interval for the HQ occurring at the 95th percentile of the population would be obtained from a large distribution of values representing Type B uncertainty for the true but unknown distribution of HQ.

3. DISCUSSION

The type B probability distributions (marginal PDF_{*i*}'s) assigned to the unknown fixed quantities in the risk assessment model are most frequently dependent on the use of professional judgment in the absence of directly relevant data. Thus, subjective confidence intervals produced for most quantitative analyses of risk assessment models must be recognized as reflecting the state of knowledge of the individuals or organizations assigned to perform the assessment. Estimates of uncertainty are required if the assessment end point is a true but unknown fixed value or a fixed but unknown distribution of values. In the first case, a subjective confidence interval is produced for the unknown true value. In the second case, a subjective confidence interval is produced for the unknown true distribution of values. For both cases, Monte Carlo simulation is used. However, the second case requires the application of Monte Carlo in two dimensions, and the ranking of the variables that contribute most to the uncertainty in the model prediction will depend on the fractile or the value of interest for

the true distribution of values the model is simulating. The quantification of uncertainty for the simulation of a true but unknown distribution of values represents the state-of-the-art in assessment modeling.

ACKNOWLEDGMENTS

The authors wish to express their sincere appreciation to Steve Simon of the National Radiological Survey of the Marshall Islands for his invaluable contribution to the preparation of Figs. 1, 2, and 3. This publication was supported by Grant R23/CCR409756-01 from the Centers for Disease Control and Prevention, National Center for Environmental Health, Radiation Studies Branch. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the Centers for Disease Control and Prevention.

REFERENCES

1. K. T. Bogen, *Uncertainty in Environmental Health Risk Assessment* (Garland, New York, 1990).
2. W. J. Conover, *Practical Non-parametric Statistics* (Wiley, New York, 1980).
3. H. C. Frey, "Separating Variability and Uncertainty in Exposure Assessment Motivations and Method," *Risk Anal.* (in press).
4. J. C. Helton, "Risk, Uncertainty in Risk, and the EPA Release Limits for Radioactive Waste Disposal," *Nuclear Technology* **101**, 18-39.
5. E. Hofer, "On Some Distinctions in Uncertainty Analysis," *Methods for Treatment of Different Types of Uncertainty*, PSAC/DOC (90) 11 (OECD Nuclear Energy Agency, Paris, 1990).
6. IAEA (International Atomic Energy Agency), "Evaluating the Reliability of Predictions Using Environmental Transfer Models," *Safety Practice Publications of the IAEA. IAEA Safety Series No. 100:1-106* (STI/PUB/835. IAEA, Vienna, Austria, 1989).
7. S. Kaplan and B. J. Garrick, "On the Quantitative Definition of Risk," *Risk Analysis* **1**(1), 11-27 (1981).
8. USEPA (United States Environmental Protection Agency), "Interim Final: Risk Assessment Guidance for Superfund. Vol I. Human Health Evaluation Manual, OSWER Directive 9285.7-01a." (EPA Office of Emergency and Remedial Response, Washington, DC, 1989).
9. USEPA, "Guidelines for Exposure Assessment," *Federal Register* **57**, No. 104, May 29 (1992).