2. VARIABILITY AND UNCERTAINTY
- 2.1. VARIABILITY VERSUS UNCERTAINTY
- 2.2. TYPES OF VARIABILITY
- 2.3. CONFRONTING VARIABILITY
- 2.4. CONCERN ABOUT UNCERTAINTY
- 2.5. TYPES OF UNCERTAINTY AND REDUCING UNCERTAINTY
- 2.6. ANALYZING VARIABILITY AND UNCERTAINTY
- 2.7. PRESENTING RESULTS OF VARIABILITY AND UNCERTAINTY ANALYSIS

REFERENCES FOR CHAPTER 2

Table 2-1. Four Strategies for Confronting Variability
Table 2-2. Three Types of Uncertainty and Associated Sources and Examples
Table 2-3. Approaches to Quantitative Analysis of Uncertainty

2. VARIABILITY AND UNCERTAINTY

The chapters that follow will discuss exposure factors and algorithms for estimating exposure. Exposure factor values can be used to obtain a range of exposure estimates such as average, high-end and bounding estimates. It is instructive here to return to the general equation for potential Average Daily Dose (ADD_{pot}) that was introduced in the opening chapter of this handbook:

$$ADD_{pot} = \frac{\text{Contaminant Concentration x Intake Rate x Exposure Duration}}{\text{Body Weight x Averaging Time}} \qquad \text{(Eqn. 2-1)}$$

With the exception of the contaminant concentration, all parameters in the above equation are considered exposure factors and, thus, are treated in fair detail in other chapters of this handbook. Each of the exposure factors involves humans, either in terms of their characteristics (e.g., body weight) or behaviors (e.g., amount of time spent in a specific location, which affects exposure duration). While the topics of variability and uncertainty apply equally to contaminant concentrations and the rest of the exposure factors in equation 2-1, the focus of this chapter is on variability and uncertainty as they relate to exposure factors. Consequently, examples provided in this chapter relate primarily to exposure factors, although contaminant concentrations may be used when they better illustrate the point under discussion.

This chapter also is intended to acquaint the exposure assessor with some of the fundamental concepts and precepts related to variability and uncertainty, together with methods and considerations for evaluating and presenting the uncertainty associated with exposure estimates. Subsequent sections in this chapter are devoted to the following topics:

- Distinction between variability and uncertainty;
- Types of variability;
- Methods of confronting variability;
- Types of uncertainty and reducing uncertainty;
- Analysis of variability and uncertainty; and
- Presenting results of variability/uncertainty analysis.

Fairly extensive treatises on the topic of uncertainty have been provided, for example, by Morgan and Henrion (1990), the National Research Council (NRC, 1994) and, to a lesser extent, the U.S. EPA (1992; 1995). The topic commonly has been treated as it relates to the overall process of conducting risk assessments; because exposure

assessment is a component of risk-assessment process, the general concepts apply equally to the exposure-assessment component.

2.1. VARIABILITY VERSUS UNCERTAINTY

While some authors have treated variability as a specific type or component of uncertainty, the U.S. EPA (1995) has advised the risk assessor (and, by analogy, the exposure assessor) to distinguish between variability and uncertainty. Uncertainty represents a lack of knowledge about factors affecting exposure or risk, whereas variability arises from true heterogeneity across people, places or time. In other words, uncertainty can lead to inaccurate or biased estimates, whereas variability can affect the precision of the estimates and the degree to which they can be generalized. Most of the data presented in this handbook concerns variability.

Variability and uncertainty can complement or confound one another. An instructive analogy has been drawn by the National Research Council (NRC, 1994: Chapter 10), based on the objective of estimating the distance between the earth and the moon. Prior to fairly recent technology developments, it was difficult to make accurate measurements of this distance, resulting in measurement uncertainty. Because the moon's orbit is elliptical, the distance is a variable quantity. If only a few measurements were to be taken without knowledge of the elliptical pattern, then either of the following incorrect conclusions might be reached:

- That the measurements were faulty, thereby ascribing to uncertainty what was actually caused by variability; or
- That the moon's orbit was random, thereby not allowing uncertainty to shed light on seemingly unexplainable differences that are in fact variable and predictable.

A more fundamental error in the above situation would be to incorrectly estimate the true distance, by assuming that a few observations were sufficient. This latter pitfall -- treating a highly variable quantity as if it were invariant or only uncertain -- is probably the most relevant to the exposure or risk assessor.

Now consider a situation that relates to exposure, such as estimating the average daily dose by one exposure route -- ingestion of contaminated drinking water. Suppose that it is possible to measure an individual's daily water consumption (and concentration of the contaminant) exactly, thereby eliminating uncertainty in the measured daily dose. The daily dose still has an inherent day-to-day variability, however, due to changes in the individual's daily water intake or the contaminant concentration in water.

It is impractical to measure the individual's dose every day. For this reason, the exposure assessor may estimate the average daily dose (ADD) based on a finite number

of measurements, in an attempt to "average out" the day-to-day variability. The individual has a true (but unknown) ADD, which has now been estimated based on a sample of measurements. Because the individual's true average is unknown, it is uncertain how close the estimate is to the true value. Thus, the variability across daily doses has been translated into uncertainty in the ADD. Although the individual's true ADD has no variability, the estimate of the ADD has some uncertainty.

The above discussion pertains to the ADD for one person. Now consider a distribution of ADDs across individuals in a defined population (e.g., the general U.S. population). In this case, variability refers to the range and distribution of ADDs across individuals in the population. By comparison, uncertainty refers to the exposure assessor's state of knowledge about that distribution, or about parameters describing the distribution (e.g., mean, standard deviation, general shape, various percentiles).

As noted by the National Research Council (NRC, 1994), the realms of variability and uncertainty have fundamentally different ramifications for science and judgment. For example, uncertainty may force decision-makers to judge how probable it is that exposures have been overestimated or underestimated for every member of the exposed population, whereas variability forces them to cope with the certainty that different individuals are subject to exposures both above and below any of the exposure levels chosen as a reference point.

2.2. TYPES OF VARIABILITY

Variability in exposure is related to an individual's location, activity, and behavior or preferences at a particular point in time, as well as pollutant emission rates and physical/chemical processes that affect concentrations in various media (e.g., air, soil, food and water). The variations in pollutant-specific emissions or processes, and in individual locations, activities or behaviors, are not necessarily independent of one another. For example, both personal activities and pollutant concentrations at a specific location might vary in response to weather conditions, or between weekdays and weekends.

At a more fundamental level, three types of variability can be distinguished:

- Variability across locations (Spatial Variability);
- Variability over time (Temporal Variability); and
- Variability among individuals (Inter-individual Variability).

Spatial variability can occur both at regional (macroscale) and local (microscale) levels. For example, fish intake rates can vary depending on the region of the country.

Higher consumption may occur among populations located near large bodies of water such as the Great Lakes or coastal areas. As another example, outdoor pollutant levels can be affected at the regional level by industrial activities and at the local level by activities of individuals. In general, higher exposures tend to be associated with closer proximity to the pollutant source, whether it be an industrial plant or related to a personal activity such as showering or gardening. In the context of exposure to airborne pollutants, the concept of a "microenvironment" has been introduced (Duan, 1982) to denote a specific locality (e.g., a residential lot or a room in a specific building) where the airborne concentration can be treated as homogeneous (i.e., invariant) at a particular point in time.

Temporal variability refers to variations over time, whether long- or short-term. Seasonal fluctuations in weather, pesticide applications, use of woodburning appliances and fraction of time spent outdoors are examples of longer-term variability. Examples of shorter-term variability are differences in industrial or personal activities on weekdays versus weekends or at different times of the day.

Inter-individual variability can be either of two types: (1) human characteristics such as age or body weight, and (2) human behaviors such as location and activity patterns. Each of these variabilities, in turn, may be related to several underlying phenomena that vary. For example, the natural variability in human weight is due to a combination of genetic, nutritional, and other lifestyle or environmental factors. Variability arising from independent factors that combine multiplicatively generally will lead to an approximately lognormal distribution across the population, or across spatial/temporal dimensions.

2.3 . CONFRONTING VARIABILITY

According to the National Research Council (NRC 1994), variability can be confronted in four basic ways (Table 2-1) when dealing with science-policy questions surrounding issues such as exposure or risk assessment. The first is to **ignore the variability** and hope for the best. This strategy tends to work best when the variability is relatively small. For example, the assumption that all adults weigh 70 kg is likely to be correct within ±25% for most adults.

The second strategy involves **disaggregating the variability** in some explicit way, in order to better understand it or reduce it. Mathematical models are appropriate in some cases, as in fitting a sine wave to the annual outdoor concentration cycle for a particular pollutant and location. In other cases, particularly those involving human characteristics or behaviors, it is easier to disaggregate the data by considering all the relevant subgroups or subpopulations. For example, distributions of body weight could be developed separately for adults, adolescents and children, and even for males and females within

each of these subgroups. Temporal and spatial analogies for this concept involve measurements on appropriate time scales and choosing appropriate subregions or microenvironments.

The third strategy is to **use the average value** of a quantity that varies. Although this strategy might appear as tantamount to ignoring variability, it needs to be based on a decision that the average value can be estimated reliably in light of the variability (e.g., when the variability is known to be relatively small, as in the case of adult body weight).

The fourth strategy involves **using the maximum or minimum value** for an exposure factor. In this case, the variability is characterized by the range between the extreme values and a measure of central tendency. This is perhaps the most common method of dealing with variability in exposure or risk assessment -- to focus on one time period (e.g., the period of peak exposure), one spatial region (e.g., in close proximity to the pollutant source of concern), or one subpopulation (e.g., exercising asthmatics). As noted by the U.S. EPA (1992), when an exposure assessor develops estimates of high-end individual exposure and dose, care must be taken not to set all factors to values that maximize exposure or dose -- such an approach will almost always lead to an overestimate.

2.4. CONCERN ABOUT UNCERTAINTY

Why should the exposure assessor be concerned with uncertainty? As noted by the U.S. EPA (1992), exposure assessment can involve a broad array of information sources and analysis techniques. Even in situations where actual exposure-related measurements exist, assumptions or inferences will still be required because data are not likely to be available for all aspects of the exposure assessment. Moreover, the data that are available may be of questionable or unknown quality. Thus, exposure assessors have a responsibility to present not just numbers, but also a clear and explicit explanation of the implications and limitations of their analyses.

Morgan and Henrion (1990) provide an argument by analogy. When scientists report quantities that they have measured, they are expected to routinely report an estimate of the probable error associated with such measurements. Because uncertainties inherent in policy analysis (of which exposure assessment is a part) tend to be even greater than those in the natural sciences, exposure assessors also should be expected to report or comment on the uncertainties associated with their estimates.

Additional reasons for addressing uncertainty in exposure or risk assessments (U.S. EPA, 1992, Morgan and Henrion, 1990) include the following:

- Uncertain information from different sources of different quality often must be combined for the assessment;
- Decisions need to be made about whether or how to expend resources to acquire additional information,;
- Biases may result in so-called "best estimates" that in actuality are not very accurate; and
- Important factors and potential sources of disagreement in a problem can be identified.

Addressing uncertainty will increase the likelihood that results of an assessment or analysis will be used in an appropriate manner. Problems rarely are solved to everyone's satisfaction, and decisions rarely are reached on the basis of a single piece of evidence. Results of prior analyses can shed light on current assessments, particularly if they are couched in the context of prevailing uncertainty at the time of analysis. Exposure assessment tends to be an iterative process, beginning with a screening-level assessment that may identify the need for more in-depth assessment. One of the primary goals of the more detailed assessment is to reduce uncertainty in estimated exposures. This objective can be achieved more efficiently if guided by presentation and discussion of factors thought to be primarily responsible for uncertainty in prior estimates.

2.5. TYPES OF UNCERTAINTY AND REDUCING UNCERTAINTY

The problem of uncertainty in exposure or risk assessment is relatively large, and can quickly become too complex for facile treatment unless it is divided into smaller and more manageable topics. One method of division (Bogen, 1990) involves classifying sources of uncertainty according to the step in the risk assessment process (hazard identification, dose-response assessment, exposure assessment or risk characterization) at which they can occur. A more abstract and generalized approach preferred by some scientists is to partition all uncertainties among the three categories of bias, randomness and true variability. These ideas are discussed later in some examples.

The U.S. EPA (1992) has classified uncertainty in exposure assessment into three broad categories:

1. Uncertainty regarding missing or incomplete information needed to fully define exposure and dose (Scenario Uncertainty).
2. Uncertainty regarding some parameter (Parameter Uncertainty).
3. Uncertainty regarding gaps in scientific theory required to make predictions on the basis of causal inferences (Model Uncertainty).

Identification of the sources of uncertainty in an exposure assessment is the first step in determining how to reduce that uncertainty. The types of uncertainty listed above can be further defined by examining their principal causes. Sources and examples for each type of uncertainty are summarized in Table 2-2.

Because uncertainty in exposure assessments is fundamentally tied to a lack of knowledge concerning important exposure factors, strategies for reducing uncertainty necessarily involve reduction or elimination of knowledge gaps. Example strategies to reduce uncertainty include (1) collection of new data using a larger sample size, an unbiased sample design, a more direct measurement method or a more appropriate target population, and (2) use of more sophisticated modeling and analysis tools.

2.6 . ANALYZING VARIABILITY AND UNCERTAINTY

Exposure assessments often are developed in a phased approach. The initial phase usually screens out the exposure scenarios or pathways that are not expected to pose much risk, to eliminate them from more detailed, resource-intensive review. Screening-level assessments typically examine exposures that would fall on or beyond the high end of the expected exposure distribution. Because screening-level analyses usually are included in the final exposure assessment, the final document may contain scenarios that differ quite markedly in sophistication, data quality, and amenability to quantitative expressions of variability or uncertainty.

According to the U.S. EPA (1992), uncertainty characterization and uncertainty assessment are two ways of describing uncertainty at different degrees of sophistication. Uncertainty characterization usually involves a qualitative discussion of the thought processes used to select or reject specific data, estimates, scenarios, etc. Uncertainty assessment is a more quantitative process that may range from simpler measures (e.g., ranges) and simpler analytical techniques (e.g., sensitivity analysis) to more complex measures and techniques. Its goal is to provide decision makers with information concerning the quality of an assessment, including the potential variability in the estimated exposures, major data gaps, and the effect that these data gaps have on the exposure estimates developed.

A distinction between variability and uncertainty was made in Section 2.1. Although the quantitative process mentioned above applies more directly to variability and the qualitative approach more so to uncertainty, there is some degree of overlap. In general, either method provides the assessor or decision-maker with insights to better evaluate the assessment in the context of available data and assumptions. The following paragraphs describe some of the more common procedures for analyzing variability and uncertainty in exposure assessments. Principles that pertain to presenting the results of variability/uncertainty analysis are discussed in the next section.

Several approaches can be used to characterize uncertainty in parameter values. When uncertainty is high, the assessor may use order-of-magnitude bounding estimates of parameter ranges (e.g., from 0.1 to 10 liters for daily water intake). Another method describes the range for each parameter including the lower and upper bounds as well as a "best estimate" (e.g., 1.4 liters per day) determined by available data or professional judgement.

When sensitivity analysis indicates that a parameter profoundly influences exposure estimates, the assessor should develop a probabilistic description of its range. If there are enough data to support their use, standard statistical methods are preferred. If the data are inadequate, expert judgment can be used to generate a subjective probabilistic representation. Such judgments should be developed in a consistent, well-documented manner. Morgan and Henrion (1990) and Rish (1988) describe techniques to solicit expert judgment.

Most approaches to quantitative analysis examine how variability and uncertainty in values of specific parameters translate into the overall uncertainty of the assessment. Details may be found in reviews such as Cox and Baybutt (1981), Whitmore (1985), Inman and Helton (1988), Seller (1987), and Rish and Marnicio (1988). These approaches can generally be described (in order of increasing complexity and data needs) as: (1) sensitivity analysis; (2) analytical uncertainty propagation; (3) probabilistic uncertainty analysis; or (4) classical statistical methods (U.S. EPA 1992). The four approaches are summarized in Table 2-3.

2.7 . PRESENTING RESULTS OF VARIABILITY AND UNCERTAINTY ANALYSIS

Comprehensive qualitative analysis and rigorous quantitative analysis are of little value for use in the decision-making process, if their results are not clearly presented. In this chapter, variability (the receipt of different levels of exposure by different individuals) has been distinguished from uncertainty (the lack of knowledge about the correct value for a specific exposure measure or estimate). Most of the data that are presented in this handbook deal with variability directly, through inclusion of statistics that pertain to the distributions for various exposure factors.

Not all approaches historically used to construct measures or estimates of exposure have attempted to distinguish between variability and uncertainty. The assessor is advised to use a variety of exposure descriptors, and where possible, the full population distribution, when presenting the results. This information will provide risk managers with a better understanding of how exposures are distributed over the population and how variability in population activities influences this distribution.

Although incomplete analysis is essentially unquantifiable as a source of uncertainty, it should not be ignored. At a minimum, the assessor should describe the rationale for excluding particular exposure scenarios; characterize the uncertainty in these decisions as high, medium, or low; and state whether they were based on data, analogy, or professional judgment. Where uncertainty is high, a sensitivity analysis can be used to credible upper limits on exposure by way of a series of "what if" questions.

Although assessors have always used descriptors to communicate the kind of scenario being addressed, the 1992 Exposure Guidelines establish clear quantitative definitions for these risk descriptors. These definitions were established to ensure that consistent terminology is used throughout the Agency. The risk descriptors defined in the Guidelines include descriptors of individual risk and population risk. Individual risk descriptors are intended to address questions dealing with risks borne by individuals within a population, including not only measures of central tendency (e.g., average or median), but also those risks at the high end of the distribution. Population risk descriptors refer to an assessment of the extent of harm to the population being addressed. It can be either an estimate of the number of cases of a particular effect that might occur in a population (or population segment), or a description of what fraction of the population receives exposures, doses, or risks greater than a specified value. The data presented in the Exposure Factors Handbook is one of the tools available to exposure assessors to construct the various risk descriptors.

However, it is not sufficient to merely present the results using different exposure descriptors. Risk managers should also be presented with an analysis of the uncertainties surrounding these descriptors. Uncertainty may be presented using simple or very sophisticated techniques, depending on the requirements of the assessment and the amount of data available. It is beyond the scope of this handbook to discuss the mechanics of uncertainty analysis in detail. At a minimum, the assessor should address uncertainty qualitatively by answering questions such as:

- What is the basis or rationale for selecting these assumptions/parameters, such as data, modeling, scientific judgment, Agency policy, "what if" considerations, etc.?

- What is the range or variability of the key parameters? How were the parameter values selected for use in the assessment? Were average, median, or upper-percentile values chosen? If other choices had been made, how would the results have differed?

- What is the assessor's confidence (including qualitative confidence aspects) in the key parameters and the overall assessment? What are the quality and the extent of the data base(s) supporting the selection of the chosen values?

Any exposure estimate developed by an assessor will have associated assumptions about the setting, chemical, population characteristics, and how contact with the chemical occurs through various exposure routes and pathways. The exposure assessor will need to examine many sources of information that bear either directly or indirectly on these components of the exposure assessment. In addition, the assessor will be required to make many decisions regarding the use of existing information in constructing scenarios and setting up the exposure equations. In presenting the scenario results, the assessor should strive for a balanced and impartial treatment of the evidence bearing on the conclusions with the key assumptions highlighted. For these key assumptions, one should cite data sources and explain any adjustments of the data.

The exposure assessor also should qualitatively describe the rationale for selection of any conceptual or mathematical models that may have been used. This discussion should address their verification and validation status, how well they represent the situation being assessed (e.g., average versus high-end estimates), and any plausible alternatives in terms of their acceptance by the scientific community.

Table 2-2 summarizes the three types of uncertainty, associated sources, and examples. Table 2-3 summarizes four approaches to analyze uncertainty quantitatively. These are described further in the 1992 Exposure Guidelines.

Table 2-1. Four Strategies for Confronting Variability		
Strategy	Example	Comment
Ignore variability	Assume that all adults weigh 70 kg	Works best when variability is small
Disaggregate the variability	Develop distributions of body weight for age/gender groups	Variability will be smaller in each group
Use the average value	Use average body weight for adults	Can the average be estimated reliably given what is known about the variability?
Use a maximum or minimum value	Use a lower-end value from the weight distribution	Conservative approach -- can lead to unrealistically high exposure estimate if taken for all factors

Table 2-2. Three Types of Uncertainty and Associated Sources and Examples

Type of Uncertainty	Sources	Examples
Scenario Uncertainty	Descriptive errors	Incorrect or insufficient information
	Aggregation errors	Spatial or temporal approximations
	Judgment errors	Selection of an incorrect model
	Incomplete analysis	Overlooking an important pathway
Parameter Uncertainty	Measurement errors	Imprecise or biased measurements
	Sampling errors	Small or unrepresentative samples
	Variability	In time, space or activities
	Surrogate data	Structurally-related chemicals
Model Uncertainty	Relationship errors	Incorrect inference on the basis for correlations
	Modeling errors	Excluding relevant variables

Table 2-3. Approaches to Quantitative Analysis of Uncertainty

Approach	Description	Example
Sensitivity Analysis	Changing one input variable at a time while leaving others constant, to examine effect on output	Fix each input at lower (then upper) bound while holding others at nominal values (e.g., medians)
Analytical Uncertainty Propagation	Examining how uncertainty in individual parameters affects the overall uncertainty of the exposure assessment	Analytically or numerically obtain a partial derivative of the exposure equation with respect to each input parameter
Probabilistic Uncertainty Analysis	Varying each of the input variables over various values of their respective probability distributions	Assign probability density function to each parameter; randomly sample values from each distribution and insert them in the exposure equation (Monte Carlo)
Classical Statistical Methods	Estimating the population exposure distribution directly, based on measured values from a representative sample	Compute confidence interval estimates for various percentiles of the exposure distribution

REFERENCES FOR CHAPTER 2

Bogen, K.T. (1990) Uncertainty in environmental health risk assessment. Garland Publishing, New York, NY.

Cox, D.C.; Baybutt, P.C. (1981) Methods for uncertainty analysis. A comparative survey. Risk Anal. 1(4):251-258.

Duan, N. (1982) Microenvironment types: A model for human exposure to air pollution. Environ. Intl. 8:305-309.

Inman, R.L.; Helton, J.C. (1988) An investigation of uncertainty and sensitivity analysis techniques for computer models. Risk Anal. 8(1):71-91.

Morgan, M.G.; Henrion, M. (1990) Uncertainty: A guide to dealing with uncertainty in quantitative risk and policy analysis. Cambridge University Press, New York, NY.

National Research Council (NRC). (1994) Science and judgment in risk assessment. National Academy Press, Washington, DC.

Rish, W.R. (1988) Approach to uncertainty in risk analysis. Oak Ridge National Laboratory. ORNL/TM-10746.

Rish, W.R.; Marnicio, R.J. (1988) Review of studies related to uncertainty in risk analysis. Oak Ridge National Laboratory. ORNL/TM-10776.

Seller, F.A. (1987) Error propagation for large errors. Risk Anal. 7(4):509-518.

U.S. EPA (1992) Guidelines for exposure assessment. Washington, DC: Office of Research and Development, Office of Health and Environmental Assessment. EPA/600/2-92/001.

U.S. EPA (1995) Guidance for risk characterization. Science Policy Council, Washington, DC.

Whitmore, R.W. (1985) Methodology for characterization of uncertainty in exposure assessments. EPA/600/8-86/009.