Technical Topics:

Uncertainty in Analyzing Health Risks

What does "uncertainty" mean?

Uncertainty is defined as "the state of being in doubt; not having sure knowledge" (from Webster's II: New Riverside Dictionary, 1988). We are often faced with uncertainty and, where numbers are involved, we use various methods of estimation. For example, a mileage scale is provided on a road map to estimate distances; the more bends in the road, the less certain is our estimate.

Risks to human health cannot be measured but can be estimated by various means. A risk estimate represents a statistical probability or chance that a health effect will occur under given circumstances. Because we do not know precisely what health effects will result from exposure to hazardous substances, our estimates of risk will always be somewhat uncertain.

An Everyday Example of Uncertainty

Suppose you are given a yardstick with no markings on it and are asked to measure the width of a room. Laying the yardstick end-to-end, you find that the room is more than five yards wide but less than six yards wide. The additional fraction appears to be about one-half the length of your yardstick. You could then estimate that the room is five-and-one-half yards wide, give or take a quarter of a yard. This is a way of making a "best estimate" and stating the range of uncertainty associated with your estimate.

Why does uncertainty exist in analyzing health risks?

Uncertainty exists in analyzing health risks primarily because we do not have complete information and because the methods of analysis are not perfect.

Technical Topics: This series of papers explains the research design, methods and terminology used in the State of Colorado's health studies related to the Rocky Flats Plant. For information about this ongoing research to identify past contaminant releases from the plant and assess potential health risks, call the Colorado Department of Health at (303) 692-2640 or 692-2652.

COLORADO DEPARTMENT OF HEALTH

In some cases, health effects occurring in specific circumstances can be calculated because a large amount of actual data exists. In 1988, for example, the U.S. Federal Highway Administration reported that 47,093 people died in vehicle accidents, resulting in a fatality rate of 2.32 deaths per one hundred million miles of vehicle travel. This rate could be used to predict an average risk of death per miles traveled for the general population. If accident data are collected for several years and used in the calculation, we would feel more confident about the risk estimate, even though the actual number of fatalities will vary from year to year.

In assessing health risks from past releases and exposures to contaminants, we seldom have complete information about individuals' actual exposures and the resulting health effects. Estimating risks becomes even more difficult if past events were not fully documented or data on contaminant releases are incomplete. That is the case in the State of Colorado's research on the health effects related to past releases of contaminants from the Rocky Flats Plant. Consequently, the study team must estimate these unknowns using whatever information is available.

Does this mean that scientists do not really know the risks related to contaminant exposure?

Scientists do not know the precise health risk related to specific contaminant exposures, but they can develop good estimates of risk. Using available data, appropriate assumptions and accepted risk assessment methods, the analyst can develop a "best estimate" of the true, but unknown, level of risk. The estimate will improve with more complete and reliable information. On the other hand, every time an assumption is made to fill in data gaps, additional uncertainty is introduced into the risk assessment.

What are some of the factors that introduce uncertainty into the analysis of past Rocky Flats contaminant releases?

Researchers do not have complete data on past releases from the Rocky Flats Plant, what happened to the contaminants after their release and how people were exposed. For instance, the amounts of contaminants released in accidents such as the 1957 fire were not measured. As a result, assumptions must be made to reconstruct the release and assess the resulting health impacts. The study team might make an assumption about the quantity of plutonium that burned, based on estimates of the amount present in the building before the fire and the rate of combustion. Because the resulting release estimate is based on such assumptions, some degree of uncertainty exists.

Computer models are often used to estimate the amounts of chemicals and radionuclides released from the plant and to simulate their movement in the environment. In some cases, the models require information that is missing or incomplete. Wind and weather data for the actual time period and location of the release might be unavailable, so the modeler could substitute data collected for that location in the same season of several other years.

The researchers fill in such data gaps with the best available information. Ultimately, they may need to make assumptions and estimates or use representative data from similar circumstances. Those assumptions or data could result in estimates that either underestimate or overestimate the contaminant releases and their potential effects.

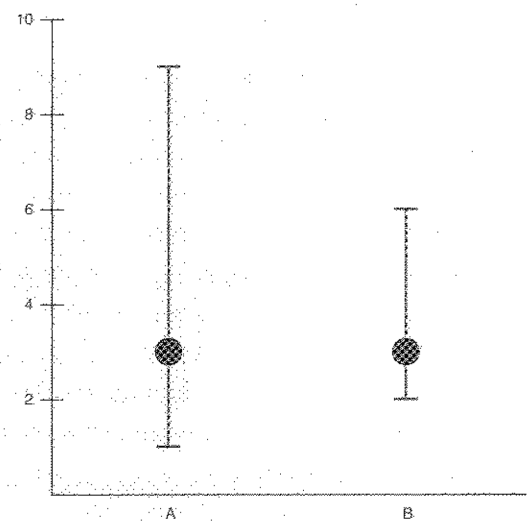
How is uncertainty expressed in risk assessments?

Uncertainty can be expressed in general statements or in numerical terms that provide a "best estimate" of risk within a range of possible estimates. Using a technique known as uncertainty analysis, the risk assessor calculates a level of confidence that the true risk value lies within the range of estimated risks. The span of this range is called a "confidence interval" and is a statistical measure of the uncertainty of the results. Uncertainty can also be expressed graphically as shown below.

On the graph, A and B represent estimates in arbitrary units of measurement. The dots show the best estimate and the vertical bars represent 90 % confidence intervals.

For A, the best estimate is 3, with 90 % confidence that the true number lies between 1.5 and 9. For B, the best estimate is also 3, with 90 % confidence that the true number lies between 2 and 6.

In both cases, the best estimate is 3, but the larger confidence interval indicates more uncertainty associated with estimate A. For B, the smaller confidence interval means less uncertainty or greater confidence. Stated another way, we are confident that the true number falls within a smaller range for estimate B than for A.



What is uncertainty analysis and what is its purpose?

Uncertainty analysis is the evaluation of the causes and extent of the data gaps in a risk assessment and their effect on the resulting risk estimates. An uncertainty analysis can be a discussion of these data gaps and their impact on the results (qualitative), or it can be a more complex mathematical evaluation of the statistical probabilities associated with a range of possible results (quantitative).

The purpose of analyzing uncertainties is to identify the largest and most important factors that affect our confidence in the results of a risk assessment. By identifying the most critical data gaps, research efforts can be focused on better defining those areas that introduce the most uncertainty into the assessment.

Can uncertainty be larger for some estimates than for others?

Yes. In some cases, the data are fairly complete and the assumptions and methods used to assess risks have a well-defined scientific basis. A relatively small amount of uncertainty, indicated by a smaller confidence interval, is associated with these assumptions (as shown in the graphic). In other cases, the basis for certain assumptions is limited, so the resulting uncertainty is higher, and the confidence interval is larger.

In general, low uncertainty is indicated by a range of estimates in which the upper and lower ends of the range are less than three times above and below the best estimate. Moderate uncertainty is indicated by a range that is greater than three but less than ten times above and below the best estimate. High uncertainty is associated with ranges where the upper and lower bounds exceed the best estimate by more than ten times.

Can anything be done to reduce uncertainty in the Rocky Flats health studies?

Uncertainty in analyzing past Rocky Flats releases and health effects cannot be eliminated because historical information is incomplete. The first step in reducing uncertainty is searching for and analyzing all possible sources of information that may be useful in the analysis. Secondly, the researchers can look for sources of supplemental data. For example, uncertainty in the data available from recent environmental sampling might be reduced by taking additional samples for comparison and verification of the previous results. Another approach to reducing uncertainty is to use different methods of analysis to fine-tune the results.

Although uncertainty cannot be eliminated, the overall goal for the Rocky Flats health studies is to perform the best possible analysis using all available relevant information to improve confidence in the results.