SECTION 3. EXPOSURE ASSESSMENT AND RISK DESCRIPTORS

The results of a risk assessment are usually communicated to the risk manager in the risk characterization portion of the assessment. This communication is often accomplished through risk descriptors which convey information and answer questions about risk, each descriptor providing different information and insights. Exposure assessment plays a key role in developing these risk descriptors, since each descriptor is based in part on the exposure distribution within the population of interest. The Risk Assessment Council (RAC) has been discussing the use of risk descriptors from time to time over the past two years.

The recent RAC efforts have laid the foundation for the discussion to follow. First, as a result of a discussion paper on the comparability of risk assessments across the Agency programs, the RAC discussed how the program presentations of risk led to ambiguity when risk assessments were compared across programs. Because different assessments presented different descriptors of risk without always making clear what was being described, the RAC discussed the advisability of using separate descriptors for population risk, individual risk, and identification of sensitive or highly exposed population segments. The RAC also discussed the need for consistency across programs and the advisability of requiring risk assessments to provide roughly comparable information to risk managers and the public through the use of a consistent set of risk descriptors.

The following guidance outlines the different descriptors in a convenient order that should not be construed as a hierarchy of importance. These descriptors should be used to describe risk in a variety of ways for a given assessment, consistent with the assessment's purpose, the data available, and the information the risk manager needs. Use of a range of descriptors instead of a single descriptor enables Agency programs to present a picture of risk that corresponds to the range of different exposure conditions encountered for most environmental chemicals. This analysis, in turn, allows risk managers to identify populations at greater and lesser risk and to shape regulatory solutions accordingly.

EPA risk assessments will be expected to address or provide descriptions of (1) individual risk to include the central tendency and high end portions of the risk distribution, (2) important subgroups of the population such as highly exposed or highly susceptible groups or individuals, if known, and (3) population risk. Assessors may also use additional descriptors of risk as needed when these add to the clarity of the presentation. With the exception of assessments where particular descriptors clearly do not apply, some form of these three types of descriptors should be routinely developed and presented for EPA risk assessments. Furthermore, presenters of risk assessment information should be prepared to routinely answer questions by risk managers concerning these descriptors.

It is essential that presenters not only communicate the results of the assessment by addressing each of the descriptors where appropriate, but they also communicate their confidence that these results portray a reasonable picture of the actual or projected exposures. This task will usually be accomplished by highlighting the key assumptions and parameters that have the greatest impact on the results, the basis or rationale for choosing these assumptions/parameters, and the consequences of choosing other assumptions.

In order for the risk assessor to successfully develop and present the various risk descriptors, the exposure assessment must provide exposure and dose information in a form that can be combined with exposure-response or dose-response relationships to estimate risk. Although there will be differences among individuals within a population as to absorption, intake rates, susceptibility, and other variables such that a high exposure does not necessarily result in a high dose or risk, a moderate or highly positive correlation among exposure, dose, and risk is assumed in the following discussion. Since the generation of all descriptors is not appropriate in all risk assessments and the type of descriptor translates fairly directly into the type of analysis that the exposure assessor must perform, the exposure assessor needs to be aware of the ultimate goals of the assessment. The following sections discuss what type of information is necessary.

1. Information about individual exposure and risk is important to communicating the results of a risk assessment.

Individual risk descriptors are intended to address questions dealing with risks borne by individuals within a population. These questions can take the form of:

- Who are the people at the highest risk?
- What risk levels are they subjected to?
- What are they doing, where do they live, etc., that might be putting them at this higher risk?
- What is the average risk for individuals in the population of interest?

The "high end" of the risk distribution is, conceptually, above the 90th percentile of the actual (either measured or estimated) distribution. This conceptual range is not meant to precisely define the limits of this descriptor, but should be used by the assessor as a target range for characterizing "high end risk". Bounding estimates and worst case scenarios[1] should not be termed high end risk estimates.

> The high end risk descriptor is a plausible estimate of the individual risk for those persons at the upper end of the risk distribution. The intent of this descriptor is to convey an estimate of risk in the upper range of the distribution, but to avoid estimates which are beyond the

[1] High end estimates focus on estimates of the exposure or dose in the actual populations. "Bounding estimates," on the other hand, purposely overestimate the exposure or dose in an actual population for the purpose of developing a statement that the risk is "not greater than...." A "worst case scenario" refers to a combination of events and conditions such that, taken together, produces the highest conceivable risk. Although it is possible that such an exposure, dose, or sensitivity combination might occur in a given population of interest, the probability of an individual receiving this combination of events and conditions is usually small, and often so small that such a combination will not occur in a particular, actual population.

> true distribution. Conceptually, high end risk means risks above about the 90th percentile of the population distribution, but not higher than the individual in the population who has the highest risk.

This descriptor is intended to estimate the risks that are expected to occur in small but definable "high end" segments of the subject population. The individuals with these risks may be members of a special population segment or individuals in the general population who are highly exposed because of the inherent stochastic nature of the factors which give rise to exposure. Where no particular difference in sensitivity can be identified within the population, the high end risk will be related to the high end exposure or dose.

In those few cases where the complete data on the population distributions of exposures and doses are available, high end exposure or dose estimates can be represented by reporting exposures or doses at selected percentiles of the distributions, such as the 90th, 95th, or 98th percentile. High end exposures or doses, as appropriate, can then be used to calculate high end risk estimates.

In the majority of cases where the complete distributions are not available, several methods help estimate a high end exposure or dose. If sufficient information about the variability in lifestyles and other factors are available to simulate the distribution through the use of appropriate modeling, e.g., Monte Carlo simulation, the estimate from the simulated distribution may be used. As in the method above, the risk manager should be told where in the high end range the

estimate is being made by stating the percentile or the number of persons above this estimate. The assessor and risk manager should be aware, however, that unless a great deal is known about exposures and doses at the high end of the distribution, these estimates will involve considerable uncertainty which the exposure assessor will need to describe.

If only limited information on the distribution of the exposure or dose factors is available, the assessor should approach estimating the high end by identifying the most sensitive parameters and using maximum or near-maximum values for one or a few of these variables, leaving others at their mean values[2]. In doing this, the exposure assessor needs to avoid combinations of parameter values that are inconsistent, e.g., low body weight used in combination with high intake rates, and must keep in mind the ultimate objective of being within the distribution of actual expected exposures and doses, and not beyond it.

If almost no data are available on the ranges for the various parameters, it will be difficult to estimate exposures or doses in the high end with much confidence, and to develop the high end risk estimate. One method that has been used in these cases is to start with a bounding estimate and "back off" the limits used until the combination of parameter values is, in the

---

[2] Maximizing all variables will in virtually all cases result in an estimate that is above the actual values seen in the population. When the principal parameters of the dose equation (e.g., concentration, intake rate, duration) are broken out into subcomponents, it may be necessary to use maximum values for more than two of these subcomponent parameters, depending on a sensitivity analysis.

judgment of the assessor, clearly within the distribution of expected exposure, and still lies within the upper 10% of persons exposed. Obviously, this method results in a large uncertainty and requires explanation.

> The risk descriptor addressing central tendency may be either the arithmetic mean risk (Average Estimate) or the median risk (Median Estimate), either of which should be clearly labeled. Where both the arithmetic mean and the median are available but they differ substantially, it is helpful to present both.

The Average Estimate, used to approximate the arithmetic mean, can be derived by using average values for all the exposure factors. It does not necessarily represent a particular individual on the distribution. The Average Estimate is not very meaningful when exposure across a population varies by several orders of magnitude or when the population has been truncated, e.g., at some prescribed distance from a point source.

Because of the skewness of typical exposure profiles, the arithmetic mean is not necessarily a good indicator of the midpoint (median, 50th percentile) of a distribution. A Median Estimate, e.g., geometric mean, is usually a valuable descriptor for this type of distribution, since half the population will be above and half below this value.

2. Information about population exposure leads to another important way to describe risk.

Population risk refers to an assessment of the extent of harm for the population as a whole. In theory, it can be calculated by summing the individual risks for all individuals within the subject population. This task, of course, requires a great deal more information than is normally, if ever, available.

Some questions addressed by descriptors of population risk include:

- How many cases of a particular health effect might be probabilistically estimated in this population for a specific time period?
- For noncarcinogens, what portion of the population are within a specified range of some benchmark level, e.g., exceedance of the RfD (a dose), the RfC (a concentration), or other health concern level?
- For carcinogens, how many persons are above a certain risk level such as 10^{-6} or a series of risk levels such as 10^{-5}, 10^{-4}, etc?

Answering these questions requires some knowledge of the exposure frequency distribution in the population. In particular, addressing the second and third questions may require graphing the risk distribution. These questions can lead to two different descriptors of population risk.

- **The first descriptor is the probabilistic number of health effect cases estimated in the population of interest over a specified time period.**

This descriptor can be obtained either by (a) summing the individual risks over all the individuals in the population when such information is available, or (b) through the use of a risk model such as carcinogenic models or procedures which assume a

linear non-threshold response to exposure. If risk varies linearly with exposure, knowing the mean risk and the population size can lead to an estimate of the extent of harm for the population as a whole, excluding sensitive subgroups for which a different dose-response curve needs to be used.

Obviously, the more information one has, the more certain the estimate of this risk descriptor, but inherent uncertainties in risk assessment methodology place limitations on the accuracy of the estimate. With the current state of the science, explicit steps should be taken to assure that this descriptor is not confused with an actuarial prediction of cases in the population (which is a statistical prediction based on a great deal of empirical data).

Although estimating population risk by calculating a mean individual risk and multiplying by the population size is sometimes appropriate for carcinogen assessments using linear, non-threshold models[3], this is not appropriate for non-carcinogenic effects or for other types of cancer models. For non-linear cancer models, an estimate of population risk must be calculated by summing individual risks. For non-cancer effects, we generally have not developed the risk assessment techniques to the point of knowing how to add risk probabilities, so a second descriptor, below, is more appropriate.

> Another descriptor of population risk is an estimate of the percentage of the population, or the number of persons, above a specified level of

---

[3] Certain important cautions apply. These cautions are more explicitly spelled out in the Agency's Guidelines for Exposure Assessment, tentatively scheduled to be published in late 1991.

risk or within a specified range of some benchmark level, e.g., exceedance of the RfD or the RfC, LOAEL, or other specific level of interest.

This descriptor must be obtained through measuring or simulating the population distribution.

3. Information about the distribution of exposure and risk for different subgroups of the population are important components of a risk assessment.

A risk manager might also ask questions about the distribution of the risk burden among various segments of the subject population such as the following:

- How do exposure and risk impact various subgroups?
- What is the population risk of a particular subgroup?

Questions about the distribution of exposure and risk among such population segments require additional risk descriptors.

> Highly exposed subgroups can be identified, and where possible, characterized and the magnitude of risk quantified. This descriptor is useful when there is (or is expected to be) a subgroup experiencing significantly different exposures or doses from that of the larger population.

These subpopulations may be identified by age, sex, life-style, economic factors, or other demographic variables. For example, toddlers who play in contaminated soil and certain high fish consumers represent subpopulations that may have greater exposures to certain agents.

> Highly susceptible subgroups can also be identified, and if possible, characterized and the magnitude of risk quantified. This descriptor is useful when the sensitivity or susceptibility to the effect for specific subgroups is (or is expected to be) significantly different from that of the larger population. In order to calculate risk for these subgroups, it will sometimes be necessary to use a different dose-response relationship.

For example, upon exposure to a chemical, pregnant women, elderly people, children, and people with certain illnesses may each be more sensitive than the population as a whole.

Generally, selection of the population segments is a matter of either *a priori* interest in the subgroup, in which case the risk assessor and risk manager can jointly agree on which subgroups to highlight, or a matter of discovery of a sensitive or highly exposed subgroup during the assessment process. In either case, once identified, the subgroup can be treated as a population in itself, and characterized the same way as the larger population using the descriptors for population and individual risk.

4. Situation-specific information adds perspective on possible future events or regulatory options.

These postulated questions are normally designed to answer "what if" questions, which are either directed at low probability but possibly high consequence events or are intended to examine candidate risk management options. Such questions might take the following form:

- What if a pesticide applicator applies this pesticide without using protective equipment?
- What if this site becomes residential in the future?
- What risk level will occur if we set the standard at 100 ppb?

The assumptions made in answering these postulated questions should not be confused with the assumptions made in developing a baseline estimate of exposure or with the adjustments in parameter values made in performing a sensitivity analysis. The answers to these postulated questions do not give information about how likely the combination of values might be in the actual population or about how many (if any) persons might be subjected to the calculated exposure or risk in the real world.

> A calculation of risk based on specific hypothetical or actual combinations of factors postulated within the exposure assessment can also be useful as a risk descriptor. It is often valuable to ask and answer specific questions of the "what if" nature to add perspective to the risk assessment.

The only information the answers to these questions convey is that if conditions A, B, and C are assumed, then the resulting exposure or risk will be X, Y, or Z, respectively. The values

for X, Y, and Z are usually fairly straightforward to calculate and can be expressed as point estimates or ranges.

Each assessment may have none, one, or several of these types of descriptors. The answers **do not** directly give information about how likely that combination of values might be in the actual population, so there are some limits to the applicability of these descriptors.