http://www.nap.edu/catalog/12209.html
We ship printed books within 1 business day; personal PDFs are available immediately.

Science and Decisions: Advancing Risk Assessment

Committee on Improving Risk Analysis Approaches Used by the U.S. EPA, National Research Council
ISBN: 0-309-12047-0, 424 pages, 7 x 10, (2008)

This PDF is available from the National Academies Press at: http://www.nap.edu/catalog/12209.html

Visit the National Academies Press online, the authoritative source for all books from the National Academy of Sciences, the National Academy of Engineering, the Institute of Medicine, and the National Research Council:

- Download hundreds of free books in PDF
- Read thousands of books online for free
- Explore our innovative research tools – try the "Research Dashboard" now!
- Sign up to be notified when new books are published
- Purchase printed books and selected PDF files

Thank you for downloading this PDF. If you have comments, questions or just want more information about the books published by the National Academies Press, you may contact our customer service department toll-free at 888-624-8373, visit us online, or send an email to feedback@nap.edu.

This book plus thousands more are available at http://www.nap.edu.

Copyright © National Academy of Sciences. All rights reserved.
Unless otherwise indicated, all materials in this PDF File are copyrighted by the National Academy of Sciences. Distribution, posting, or copying is strictly prohibited without written permission of the National Academies Press. Request reprint permission for this book.

THE NATIONAL ACADEMIES
Advisers to the Nation on Science, Engineering, and Medicine

Science and Decisions: Advancing Risk Assessment
http://www.nap.edu/catalog/12209.html

5

Toward a Unified Approach to Dose-Response Assessment

THE NEED FOR AN IMPROVED DOSE-RESPONSE FRAMEWORK

Introduction to the Problem

As described in Chapter 4, one of the urgent challenges to risk assessment is the evaluation of hazard and risk in a manner that is faithful to the underlying science, is consistent among chemicals, accounts adequately for variability and uncertainty, does not impose artificial distinctions among health end points, and provides information that is maximally useful for risk characterization and risk management. There have been efforts to harmonize dose-response methods for cancer and noncancer end points, but, as discussed below, criticisms have been raised regarding the validity of dose-response assessments for risk characterizations and management and regarding the treatment of uncertainty and variability in human sensitivity. This chapter examines the science governing dose-response assessment for a variety of end points (cancer and noncancer) and develops an integrative framework that provides conceptual and methodologic approaches for cancer and noncancer assessments.

Current Framework

Dose-response assessments for carcinogenic end points have been conducted very differently from noncancer assessments. For carcinogens, it has been assumed that there is no threshold of effect, and dose-response assessments have focused on quantifying the risk at low doses. The current Environmental Protection Agency (EPA) approach derives a "point of departure" (POD), such as the lower bound on the dose that results in an excess risk of 10% based on fitting of a dose-response model to animal bioassay data (EPA 2000a). After adjustment for animal-human differences in the dose metric, risk is assumed to decrease linearly with doses below the POD for carcinogens that are direct mutagens or are associated with large human body burdens (EPA 2005a). The population burden of disease or the population risk at a given exposure is estimated. In practice, EPA carcinogen assessments do

Copyright © National Academy of Sciences. All rights reserved.

Science and Decisions: Advancing Risk Assessment
http://www.nap.edu/catalog/12209.html

not account for differences among humans in cancer susceptibility other than from possible early-life susceptibility (see Chapter 4).

For noncancer end points, it is assumed that homeostatic and defense mechanisms lead to a dose threshold[1] (that is, there is low-dose nonlinearity), below which effects do not occur or are extremely unlikely. For these agents, risk assessments have focused on defining the reference dose (RfD) or reference concentration (RfC), a putative quantity that is "likely to be without an appreciable risk of deleterious effects" (EPA 2002a, p. 4-4). The "hazard quotient" (the ratio of the environmental exposure to the RfD or RfC) and the "hazard index" (HI, the sum of hazard quotients of chemicals to which a person is exposed that affect the same target organ or operate by the same mechanism of action) (EPA 2000b) are sometimes used as indicators of the likelihood of harm. An HI less than unity is generally understood as being indicative of lack of appreciable risk, and a value over unity indicates some increased risk. The larger the HI, the greater the risk, but the index is not related to the likelihood of adverse effect except in qualitative terms: "the HI cannot be translated to a probability that adverse effects will occur, and is not likely to be proportional to risk" (EPA 2006a). Thus, current RfD-based risk characterizations do not provide information on the fraction of the population adversely affected by a given dose or on any other direct measure of risk (EPA 2000a). That deficiency is present whether the dose is above the RfD (in which case the risk may be treated as nonzero but is not quantified) or below the RfD (in which case the risk can be treated as "unappreciable" or zero even though with some unquantified probability it is not zero).

As in cancer dose-response assessment, the RfD is also derived from a POD, which could be a no-observed-adverse-effect level (NOAEL) or a benchmark dose (BMD). However, instead of extrapolating to a low-dose risk, the POD is divided by "uncertainty factors" to adjust for animal-human differences, human-human differences in susceptibility, and other factors (for example, data gaps or study duration). In a variant of the RfD approach to noncancer or low-dose nonlinear cancer risk assessment, the agency calculates a "margin of exposure" (MOE), the ratio of a NOAEL or POD to a projected environmental exposure (EPA 2000a, 2005b). The MOE is compared with the product of uncertainty factors; an MOE greater than the product is considered to be without appreciable risk or "of low concern," and an MOE smaller than the product reflects a potential health concern (EPA 2000b). MOEs and RfDs are defined for durations of exposure (for example, acute, subchronic, and chronic) and may be defined for specific life stages (for example, developmental) (EPA 2002a).

Recent refinements in risk-assessment methods in EPA have used mode-of-action (MOA)[2] evaluations in dose-response assessment. EPA's *Guidelines for Carcinogen Risk Assessment* (2005b) state that if a compound is determined to be "DNA reactive and [to] have direct mutagenic activity" or to have high human exposures or body burdens "near doses associated with key precursor events" (EPA 2005b, p. 3-21), a no-threshold approach is applied; risk below the POD is assumed to decrease linearly with dose. For carcinogens with sufficient MOA data to conclude nonlinearity at low doses, such as those acting through a cytotoxic MOA, the RfD approach outlined above for noncancer end points is applied (EPA 2005b),

[1]More recent noncancer guidelines have abandoned the term *threshold*, noting the difficulty of empirically distinguishing dose-response relationships with true biologic thresholds from ones that are nonlinear at low doses (EPA 2005b, p. 3-24).

[2]Following EPA 2005b (p. 1-10), the MOA is defined as "a sequence of key events and processes, starting with interaction of an agent with a cell, proceeding through operational and anatomical changes, and resulting" in the adverse effect. "A 'key event' is an empirically observable precursor step that is itself a necessary element of the mode of action or is a biologically based marker for such an element."

Copyright © National Academy of Sciences. All rights reserved.

Science and Decisions: Advancing Risk Assessment
http://www.nap.edu/catalog/12209.html

except when there is adequate evidence to support mechanistic modeling (there has been no such case).

Another refinement in dose-response assessment has been the derivation of the RfD or low-dose cancer risk from a POD that is calculated using BMD methodology (EPA 2000a). In noncancer risk assessment, this approach has the advantage of making better use of the dose-response evidence available from bioassays than do calculations based on NOAELs. It also provides additional quantitative insight into the risk presented in the bioassay at the POD because for quantal end points the POD is defined in terms of a given risk for the animals in the study.

EPA's treatment of noncancer and low-dose nonlinear cancer end points is a major step by the agency in an overall strategy to harmonize cancer and noncancer approaches to dose-response assessment. Other aspects of this harmonization for the different end points include consideration of the same cross-species factors (EPA 2006b), and the same pharmacokinetic adjustments. EPA staff have also explored for noncancer end points dose-response modeling that results in probabilistic descriptions (for example, for acrolein, Woodruff et al. 2007) and that could be readily integrated into benefits evaluation (for thyroid-disrupting chemicals, Axelrad et al. 2005). But these approaches have not found their way into agency practice.

Scientific, Technical, and Operational Problems with the Current Approach

The committee recognizes EPA's efforts to examine and refine dose-response assessment methodology and practice and the agency's work to clarify its approaches and practices in guidelines and other documents (for example, EPA 2000a, 2002b, 2004, 2005b). A number of improvements over the last decade can be noted, such as the movement toward using MOA determinations and the application of BMD methods. However, the current framework has important structural problems, some of which have been exacerbated by recent decisions. Figure 5-1 presents an outline of the current framework for dose-response assessment and risk characterization in EPA and some major limitations in the framework, which are discussed below.

Potential Low-Dose Linearity for Noncancer and "Nonlinear" Cancer End Points

Thresholds are assumed for noncarcinogens and for carcinogens believed to operate through an MOA considered nonlinear at low doses. The rationale is that at levels below the threshold dose, clearance pathways, cellular defenses, and repair processes have been thought to minimize damage so that disease does not result. However, as illustrated in Figure 5-2, threshold determinations should not be made in isolation, inasmuch as other chemical exposures and biologic factors that influence the same adverse effect can modify the dose-response relationship at low doses and should therefore be considered.

Nonlinear Cancer End Points

The current determination of "nonlinearity" based on MOA assessment is a reasonable approach to introduce scientific evidence on MOA into cancer dose-response assessment. However, some omissions in this overall approach for low-dose nonlinear carcinogens could yield inaccurate and misleading assessments. For example, the current EPA practice of determining "nonlinear" MOAs does not account for mechanistic factors that can create linearity at low dose. The dose-response relationship can be linear at a low dose when an exposure contributes to an existing disease process (Crump et al. 1976, Lutz 1990). Effects

Copyright © National Academy of Sciences. All rights reserved.

Science and Decisions: Advancing Risk Assessment
http://www.nap.edu/catalog/12209.html

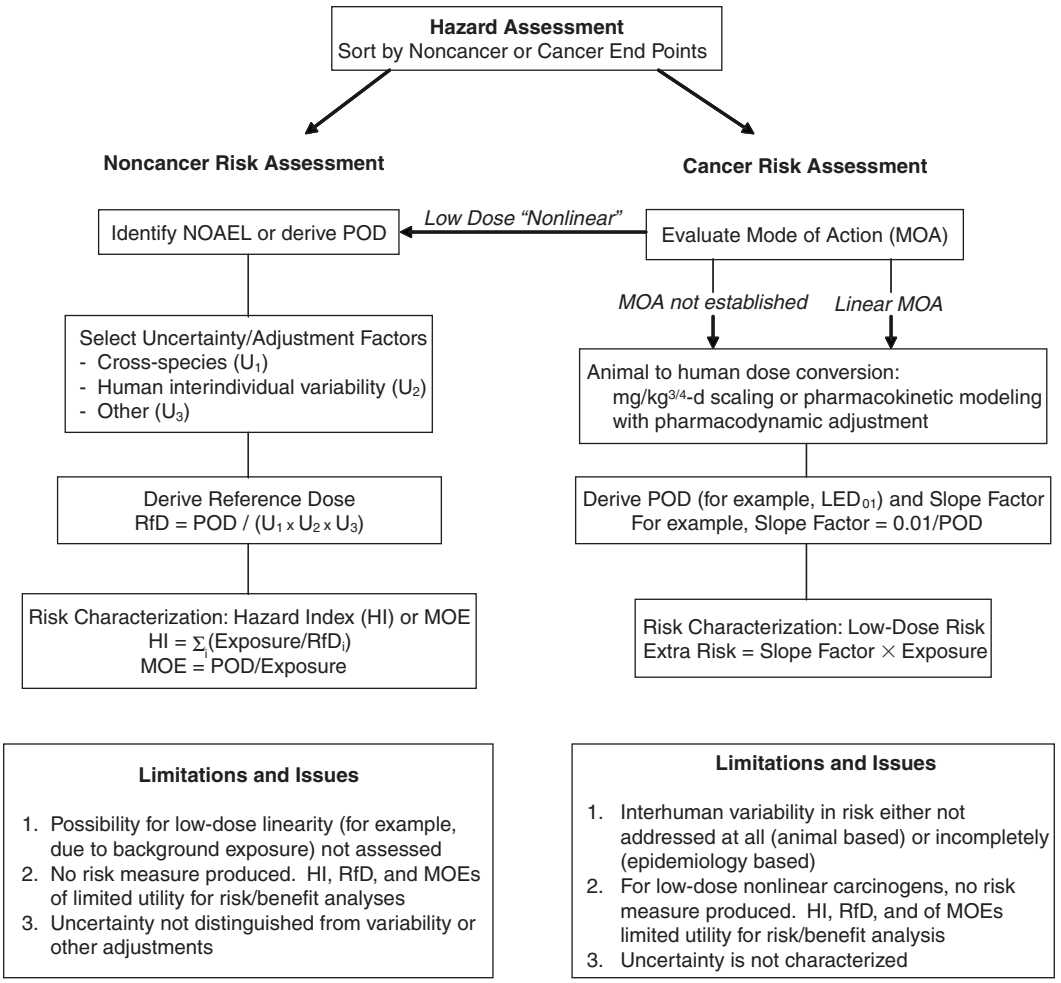

FIGURE 5-1 Current approach to noncancer and cancer dose-response assessment.

of exposures that add to background processes and background endogenous and exogenous exposures can lack a threshold if a baseline level of dysfunction occurs without the toxicant and the toxicant adds to or augments the background process. Thus, even small doses may have a relevant biologic effect. That may be difficult to measure because of background noise in the system but may be addressed through dose-response modeling procedures. Human variability with respect to the individual thresholds for a nongenotoxic cancer mechanism can result in linear dose-response relationships in the population (Lutz 2001).

In the laboratory, nonlinear dose-response processes—for example, cytotoxicity, impaired immune function and tumor surveillance, DNA methylation, endocrine disruption, and modulation of cell cycles—may be found to cause cancer in test animals. However, given the high prevalence of those background processes, given cancer as an end point, and given the multitude of chemical exposures and high variability in human susceptibility, the results may still be manifested as low-dose linear dose-response relationships in the human population (Lutz 2001). The possibility of low-dose linearity due to background is acknowledged

Copyright © National Academy of Sciences. All rights reserved.

Science and Decisions: Advancing Risk Assessment
http://www.nap.edu/catalog/12209.html

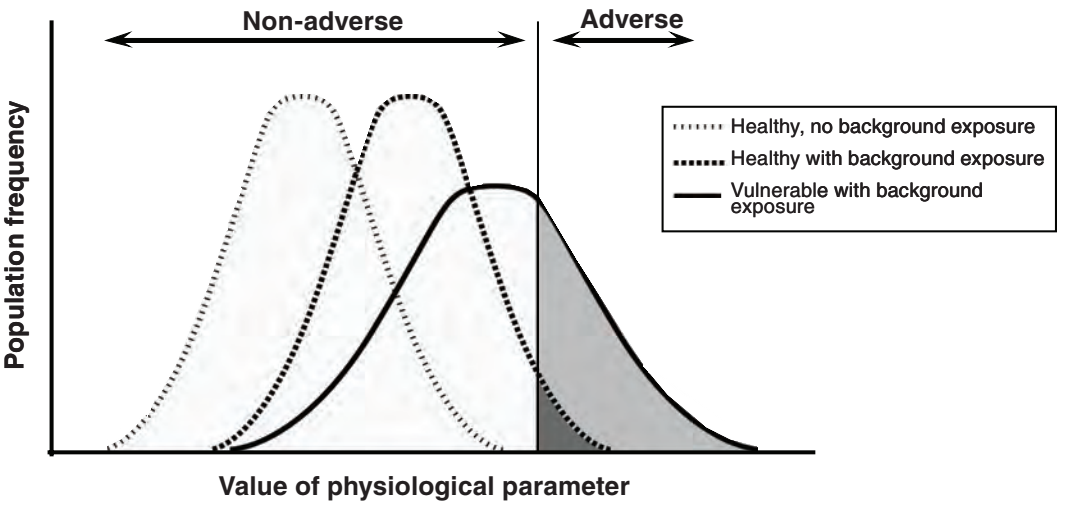

FIGURE 5-2 Value of physiologic parameter for three hypothetical populations, illustrating that population responses depend on a milieu of endogenous and exogenous exposures and on vulnerability of population due to health status and other biologic factors. Source: Adapted from Woodruff et al. 2007. Reprinted with permission; copyright 2007, *Environmental Health Perspectives*.

in the EPA (2005b) *Guidelines for Carcinogen Risk Assessment* to a limited degree—for chemicals with high body burdens or high exposures—but has not been addressed in EPA assessments. And EPA practices do not call for systematic evaluation of endogenous and exogenous exposures or mechanisms that can lead to linearity.

By segregating cancer and noncancer risk assessment, the current framework tends to place undue focus on "complete" carcinogens, ignoring contributions to ongoing carcinogenesis processes and the multifactorial nature of cancer. Chemicals that may increase human cancer risk by contributing to an underlying process are handled essentially as noncarcinogens even though they may be integral to the carcinogenic process. The dichotomy increases the burden of judging which chemicals are carcinogens rather than accepting the variety of carcinogenic MOAs and incorporating them into a comprehensive risk assessment.

Noncancer End Points

Similarly, noncarcinogens can exhibit low-dose linearity, for example, when there is considerable interindividual variability in susceptibility and each individual has his or her own threshold, especially when an underlying disease (such as cardiopulmonary disease) can interact with the toxicant (such as particulate matter [PM] or ozone). Schwartz et al. (2002) made the argument for the absence of a population threshold for mortality effects of PM. Other factors that support nonthreshold dose-response relationships for noncarcinogens include

- The observation of dose-response relationships with no apparent thresholds for subtle, common adverse end points, such as IQ loss or neurobehavioral deficits associated with lead or methylmercury exposures—an observation that continues to be made even as investigators probe for effects at smaller exposures (Axelrad et al. 2007). Those effects occur at lower doses than frank toxicity and are expected to become a more common basis of

Copyright © National Academy of Sciences. All rights reserved.

Science and Decisions: Advancing Risk Assessment
http://www.nap.edu/catalog/12209.html

dose-response assessment as increasingly subtle end points are studied with more sensitive tests (for example, tests based on -omics) or epidemiologically.

- The fact that in receptor-mediated events, even at very low doses a chemical can occupy receptor sites and theoretically perturb cell function (such as signal transduction or gene expression) or predispose the cell to other toxicants that bind to or modulate the receptor system (such as organochlorines and the aryl hydrocarbon receptor or endocrine disruptors and hormonal binding sites) (Brouwer et al. 1999; Jeong et al. 2008).
- The observation that exposures that perturb or accelerate background endogenous disease processes and add to background endogenous and exogenous exposures may not show evidence of a threshold, as described above ("Nonlinear Cancer End Points").

There are multiple toxicants (for example, PM and lead) for which low-dose linear concentration-response functions rather than thresholds have been derived for noncancer end points. The current EPA framework treats them as exceptions (implicitly if not explicitly) and does not provide methods and practices for readily assessing the dose-response relationship for cases in which thresholds are not apparent or not expected, for example, because of background additivity. As discussed in this chapter, for critical end points driving the risk characterization at low doses, such cases may be common, and a new framework and practice are needed.

Another problem posed by the current noncancer framework is that the term *uncertainty factors* is applied to the adjustments made to calculate the RfD to address species differences, human variability, data gaps, study duration, and other issues. The term engenders misunderstanding: groups unfamiliar with the underlying logic and science of RfD derivation can take it to mean that the factors are simply added on for safety or because of a lack of knowledge or confidence in the process. That may lead some to think that the true behavior of the phenomenon being described may be best reflected in the unadjusted value and that these factors create an RfD that is highly conservative. But the factors are used to adjust for differences in individual human sensitivities, for humans' generally greater sensitivity than test animals' on a milligrams-per-kilogram basis, for the fact that chemicals typically induce harm at lower doses with longer exposures, and so on. At times, the factors have been termed *safety factors*, which is especially problematic given that they cover variability and uncertainty and are not meant as a guarantee of safety.

The Need for Evaluation of Background Exposures and Predisposing Disease Processes

Dose-response assessments for noncancer and nonlinear cancer end points are generally performed without regard to exposure to other chemicals that affect the same pathologic processes or the extent of pre-existing disease in the population. The need to address chemicals that have "a common mechanism of toxicity" in a cumulative risk assessment has been established for pesticides under the Food Quality Protection Act (FQPA) of 1996 (EPA 2002b, p. 6). EPA (2002b) provides a useful example, but it was driven principally by the explicit requirements of the FQPA, and few noncarcinogens are evaluated in this way. Furthermore, dose additivity has been observed at relatively low doses for various endocrine-related toxicities with similar and dissimilar mechanisms of action (for example, Gray et al. 2001; Wolf et al. 2004; Crofton et al. 2005; Hass et al. 2007; Metzdorff et al. 2007). Dosing animals with two chemicals that have different MOAs at their NOAELs resulted in a significant adverse response, which suggested dose additivity (as when two chemicals at subthreshold doses lead to an effect). In practice, a common implicit assumption is effect

Copyright © National Academy of Sciences. All rights reserved.

Science and Decisions: Advancing Risk Assessment
http://www.nap.edu/catalog/12209.html

additivity—two subthreshold doses yield a nonresponse because neither produces a response on its own.

Consideration of chemicals that have a common MOA has not included how endogenous and other chemicals, not the direct subjects of testing and evaluation by regulatory agencies, affect the human dose-response relationship. The recent EPA draft dibutyl phthalate (DBP) assessment is an example in which there was an opportunity to consider cumulative exposure to the various agents that can contribute to the antiandrogen syndrome seen with phthalates, but the impact of even other phthalates on the DBP dose-response relationship was not taken into account in setting the draft RfD (EPA 2006c). In the application of such an assessment, DBP exposures above the RfD would be treated as posing some undefined extra degree of risk and DBP exposures below the RfD would, without further guidance from the agency, potentially be treated as riskfree without regard to the presence of other antiandrogen exposures.

Risk-Assessment Outcomes Needed for Risk Evaluation and Benefit Analysis

The end products of noncancer (and nonlinear cancer) assessments in the current paradigm (exposure-effect quotients that qualitatively indicate potential risk—MOEs, RfDs, and RfCs, Figure 5-1) are inadequate for benefit-cost analyses or for comparative risk analyses. MOEs and RfDs as currently defined do not provide a basis for formally quantifying the magnitude of harm at various exposure levels. Therefore, the committee finds the 2005 *Guidelines for Carcinogen Risk Assessment* movement toward RfDs and away from an expression of risk posed by nonlinear carcinogens problematic. Similarly, although noncancer risk assessment has moved to a BMD framework that makes better use of evidence than an approach based on NOAELs and lowest observed-adverse-effect levels (LOAELs), the paradigm remains one of defining an RfD or RfC without any sense of the degree of population risk reduction that would be found in moving from one dose to another dose. A probabilistic approach to noncancer assessment, similar to how cancer risks are expressed, would be much more useful in risk-benefit analysis and decision-making. The current threshold-nonthreshold dichotomy creates an inconsistent approach for bringing toxicology and risk science into the decision-making process.

That paradigm has other unintended consequences. For example, the linear-extrapolation exercise for carcinogens and lack of consideration of linearity for noncarcinogens and "nonlinear" carcinogens create a high bar of evidence for carcinogen identification and reduce the consideration of the possibility of noncancer end points for carcinogens. More generally, the many noncancer health end points are generally given little weight in benefit-cost analyses or other analytically driven decision frameworks in part because of the nature of the resulting qualitative risk characterization.

In the general case in which an intervention reduces exposures from above the RfD to below the RfD, it is particularly unfortunate to fail to quantify this benefit. It might be possible, through economic valuation (willingness-to-pay or contingent-valuation) studies, to estimate the benefits of moving N members of the population from exposure above the RfD to exposure below the RfD, but it would be more straightforward and intelligible to directly estimate the benefits of such an exposure and risk reduction. The current approach also does not address the benefits of lowering exposures that are already below the RfD or the benefits of lowering exposures from above the RfD to an exposure level that is still above the RfD, both of which, if understood to be associated with a nonzero probability of harm, also need valuation. The framework described below provides a means of generating the data needed for such analyses.

Copyright © National Academy of Sciences. All rights reserved.

Science and Decisions: Advancing Risk Assessment
http://www.nap.edu/catalog/12209.html

Limitations of the Current Approach for Low-Dose Linear Cancer End Points

EPA assumes that the linear default approach for dose-response assessment provides "an upper-bound calculation of potential risk at low doses," which is "thought to be public-health protective at low doses for the range of human variation" (EPA 2005b, p. A-9). EPA (2005b) noted that the National Research Council reports (NRC 1993, 1994) generally discussed the variability in human susceptibility to carcinogens and that EPA and other agencies were conducting research on the issue. The committee finds that although the precise degree of human variability is not known, the upper statistical bound derived from fits to animal data does not address human variation, as discussed below. Further, with few exceptions (EPA 2001a), the current practice embeds an implicit assumption that it is zero. This is not credible and is increasingly unwarranted as more and more studies document the substantial interindividual variation in the human population (see Chapter 4).

According to EPA, "the linear default procedure adequately accounts for human variation unless there is case-specific information for a given agent or mode of action that indicates a particularly susceptible subpopulation or lifestage, in which case the special information will be used" (EPA 2005b, p. A-9). That implies that in general the linear-extrapolation procedure will overestimate the risk to an extent that will account for the underestimation bias related to the omission of human heterogeneity. EPA provides no evidence to support that assumption and in essence establishes a default (no variability in susceptibility) that is unsubstantiated (see Chapter 6 for discussion of "missing" defaults). There are three main steps in deriving human cancer risk from animal bioassay data: adjusting animal doses to equivalent human doses, deriving the POD by fitting a mathematical model to the data, and linearly extrapolating from the POD to lower doses. The default animal-to-human adjustment is based on metabolic differences due to the roughly 200- to 2,000-fold differences in body sizes and is set at a median value without accounting for the large qualitative uncertainty, in any particular application, of the humans being more sensitive than the animal or vice versa. The lower bound on the POD merely accounts for the uncertainty in the model fitted to data from the fairly homogeneous animals used in studies. If the true dose-response relationship for an agent is indeed linear, the statistical lower confidence limit (for example, the BMD lower confidence limit [BMDL]) associated with a POD (for example, the BMD) provides a small increment of "conservatism"—typically not more than a factor of 2 (Subramaniam et al. 2006). That is highly unlikely to account for variation in susceptibility in cancer in a large exposed human population (see Chapter 4). If, instead, the true dose-response relationship is nonlinear, treating it as linear might introduce enough "conservatism" to offset the underestimation of risk in people of above-average susceptibility, but the degree to which the high-dose-based estimate is in error would preferably be analyzed separately. The practice of assuming no human variation in response to compounds for which linearity is applied is simplistic and inconsistent with the manner in which noncancer assessments are conducted. Many factors can cause the cancer response to be highly variable in the population, including age, sex, genetic polymorphisms, endogenous disease processes, lifestyle, and coexposure to other xenobiotics common in the human environment (see "Variability and Vulnerability in Risk Assessment" in Chapter 4). Some of those factors, especially pharmacokinetics and early age, are beginning to be considered in a few cancer risk assessments, but much more emphasis needs to be placed on describing the ranges of susceptibility and risk.

Other Limitations of the Current Approach

One cross-cutting issue for all end points is the degree to which dose-response characterization is done in data-poor cases. Often, a compound on which information is sparse is

Copyright © National Academy of Sciences. All rights reserved.

Science and Decisions: Advancing Risk Assessment
http://www.nap.edu/catalog/12209.html

not addressed in a quantitative risk assessment and operationally can be treated as though it posed no risk of regulatory importance. That is unlikely to describe the situation adequately or to be helpful in setting research priorities. An approach to that problem is described in Chapter 6.

In addition, any analysis must grapple with the best approach for integrating data from multiple studies and on multiple end points. There has been a tendency in risk assessment to pick a single dataset with which to describe risk, in part because it leads to straightforward rationales that are easy to explain, understand, and communicate. However, the direction toward better understanding of uncertainty, human variability, and more accurate assessment necessarily involves increasing complexity and integration of evidence from disparate sources. It also may involve constructing dose-response relationships based on evidence from a variety of study types (such as cancer bioassays and in vitro studies). Also, a given exposure to a particular chemical may affect multiple end points, and a risk description based on one tumor site or effect may fall short of conveying the overall risk posed by the substance.

In summary, the committee finds multiple scientific and operational limitations in the current approach for both cancer and noncancer risk assessments. The following section describes a means for addressing many of the issues by developing a unified framework for toxicity assessment that incorporates variability and uncertainty more completely and provides quantitative risk information on cancer and noncancer end points alike.

A UNIFIED FRAMEWORK AND APPROACH FOR DOSE-RESPONSE ASSESSMENT

The committee finds that the underlying science is more consistent with a new conceptual framework for dose-response modeling and recommends that the agency adopt a unified framework. Figure 5-3a illustrates the underlying dose-response principles for the framework, which includes background processes and exposures in considering risks on the individual and population scales. Figure 5-3a shows that an individual's risk from exposure to an environmental chemical is determined by the chemical itself, by concurrent background exposures to other environmental and endogenous chemicals that affect toxicity pathways and disease processes, and by the individual's biologic susceptibility due to genetic, lifestyle, health, and other factors. How the population responds to chemical insults depends on individual responses, which vary among individuals.

Clearly, background exposures and biologic susceptibility factors differ substantially between animals and humans, and there can be more confidence in dose-response descriptions that consider and account for background exposure and biologic susceptibility of populations for which risks are being estimated. Figure 5-3b provides a depiction of individual and population risk that formally takes these factors into account. The shape of the population dose-response relationship at low doses is inferred from an understanding of individual dose-response relationships, which in turn are based on consideration of background exposure and biologic susceptibility on human heterogeneity. An upper bound on the population dose-response relationship would be derived to express uncertainty in the population dose-response relationship. For compounds whose effects show a linear dose-response relationship, this upper bound is not the same as the familiar upper bound derived by fitting dose-response models to animal bioassay data. The latter upper bound measures only a very small aspect of uncertainty: that due to sampling variability and the statistical fit to animal data. Here, the committee envisions a more comprehensive description of uncertainty that accounts for other aspects, such as uncertainty in cross-species extrapolation. The dose of the environmental chemical that poses, say, a risk above background ("extra risk") of 10^{-5} in a population, could be described by a probability distribution that reflects

Copyright © National Academy of Sciences. All rights reserved.

Science and Decisions: Advancing Risk Assessment
http://www.nap.edu/catalog/12209.html

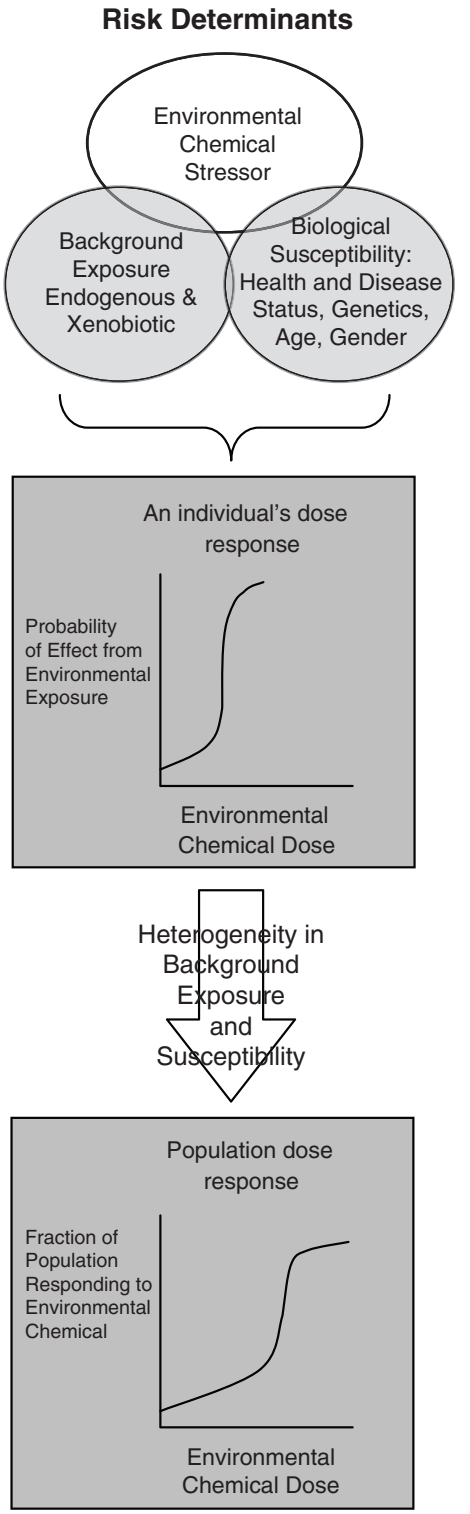

FIGURE 5-3a New conceptual framework for dose-response assessment. Risk posed by environmental chemical is determined from individual's biologic make-up, health status, and other endogenous and exogenous exposures that affect toxic process; differences among humans in these factors affect shape of population dose-response curve.

Copyright © National Academy of Sciences. All rights reserved.

Science and Decisions: Advancing Risk Assessment
http://www.nap.edu/catalog/12209.html

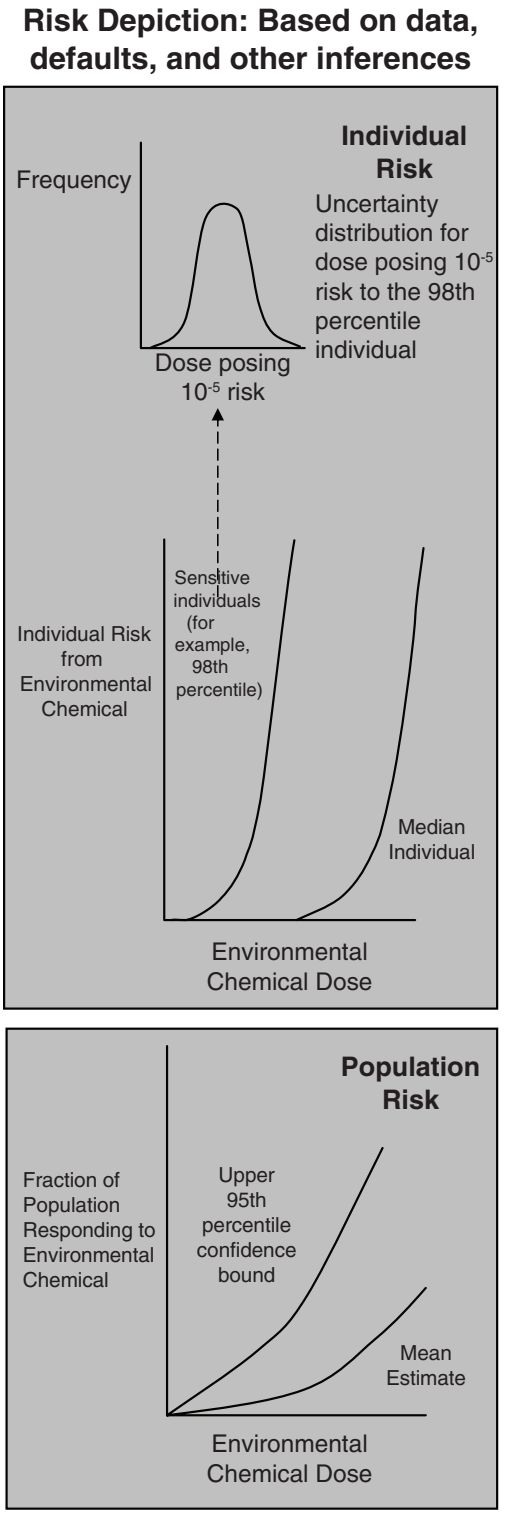

FIGURE 5-3b Risk estimation and description under the new conceptual framework for dose-response assessment. Risk estimates are based on inferences made from human, animal, MOA, and other data and understanding of possible background processes and exposures. Ideally, population dose-response relationship and uncertainty (represented by upper 95% bound) and dose-response relationships for sensitive members of population are described. (As explained in text, upper 95% confidence bound on risk is not same as upper-bound estimate generated in current cancer risk assessments.) Mean estimate of population risk can be derived from understanding of individual risk.

Copyright © National Academy of Sciences. All rights reserved.

Science and Decisions: Advancing Risk Assessment
http://www.nap.edu/catalog/12209.html

the uncertainty. Ideally, risk would be estimated for sensitive as well as typical individuals, and uncertainty in those estimates would also be described.

One important outcome of the new approach is the redefinition of the RfD as a risk-specific dose rather than as a dichotomous risk–unappreciable risk descriptor. The redefinition is described further below.

Characteristics of the Dose-Response Framework

The dose-response framework envisioned includes the following features:

- *Dose-response characterizations that use the spectrum of evidence from human, animal, mechanistic, and other relevant studies.* Whole-animal dose-response studies will continue to play a central role in establishing PODs for most chemicals, but information on human heterogeneity, background exposures, and disease processes and data from mechanistic in vitro and in vivo studies will be critical in selecting the approach to the dose-response analysis. Some information used in the dose-response derivation will be chemical-specific. In the absence of reliable chemical-specific information on human variability, interspecies differences, and other components of the analysis, generalizations and defaults based on evidence from other chemicals and end points and theoretical considerations may be used. Clearly, this presents challenges associated with selection of data sources, data synthesis, and model uncertainty.
- *The goal of providing a probabilistic characterization of harm*, such as a description of the form "at dose D, R fraction of the population would be anticipated to suffer harm with a confidence interval of R_L-R_H." For example, a summary statement of risk may be that at an air concentration of 0.05 ppm (= D), 1/10,000 (= R) of the population are likely to be affected with a 95% confidence interval (CI) of 5/100,000-3/10,000 of the population. That general form can be made more specific to particular outcomes and MOAs. For example, as described later in this chapter, for agents unlikely to have a threshold even at the individual level (such as mutagenic carcinogens), each person is assumed to be at a finite risk, and one can also make statements about individual risk. A summary statement may be given like that above with a further description that the 95th percentile individual at a dose of 0.05 ppm may face a risk of 1/1,000 (with a CI of 5/10,000-3/1,000). Thus, for a population uniformly exposed to a compound at 0.05 ppm, the characterization would indicate the distribution of risk among individuals (with variability driven by differences in background exposures and biologic susceptibility), in this example, with 5% of individuals having estimated risks above 1/1,000 (with associated confidence bounds). The key attribute of the characterization would be a quantitative and probabilistic characterization of harm for each critical end point. A similar position for probabilistic expression of noncancer risk has been advocated by the EPA Science Advisory Board (EPA SAB 2002). Multiple end points of varied severity would be considered. In many cases, new research or well-justified default approaches will be needed to attain this level of refinement in noncancer dose-response analysis.
- *Explicit consideration of human heterogeneity in response,* for both cancer and noncancer end points, that is distinguished from uncertainty. This variability assessment would consider susceptibility due to age, sex, health status, genetic makeup, and other factors. Uncertainty in human variability estimates would be described, preferably quantitatively. The rigor of this characterization would be commensurate with the needs of the assessment (see Chapters 3 and 4).
- *Treatment of uncertainty aimed at characterizing the most important types of uncertainties* for both cancer and noncancer end points. This could involve formal quantification

Copyright © National Academy of Sciences. All rights reserved.

Science and Decisions: Advancing Risk Assessment
http://www.nap.edu/catalog/12209.html

following probabilistic approaches that are consistent with recommendations about the use of default assumptions in Chapter 6. It could also include sensitivity analyses or qualitative characterizations if they would provide a better description of uncertainty or are commensurate with the needs of the assessment.

- *Evaluation of background exposure and susceptibility in order to select modeling approach*. The assessment of "background exposure" and "background disease processes" would involve characterization of other chemicals or nonchemical stressors that influence the same general pathologic processes as the chemical under evaluation. Such consideration should aid the evaluation of the shape of the dose-response relationship, including the potential for low-dose linearity and high-risk subpopulations and hence appropriate methodologic approaches for the dose-response analysis. Background exposures and susceptibility factors can result in linear low-dose-response relationships that would otherwise be considered low-dose nonlinear on the basis of MOA alone.
- *Use of distributions instead of "uncertainty factors,"* as the science and data develop and are found to provide a sufficient basis for doing so. For example, research is going on to develop uncertainty distributions for the pharmacokinetic (PK) and pharmacodynamic (PD) components of the interspecies and intraspecies human uncertainty factors (for example, Hattis and Lynch 2007). Data-driven adjustment factors developed by such bodies as the World Health Organization's International Program on Chemical Safety (IPCS 2005) are being expanded to probabilistic descriptions on the basis of information from the pharmaceutical sector and emerging from the biologic sciences. It will be a challenge to overcome some of the data limitations for developing those approaches. For example, many studies use small numbers of human subjects, so the sensitive individuals in the population may not be characterized quantitatively by distributions derived from these studies, particularly if the true human distribution is multimodal. Approaches are needed to address that issue. The formal incorporation of variability due to polymorphisms, aging, endogenous disease status, exposure, and other factors will probably prove to be complex and challenging. Later in this chapter, examples are given of an approach for developing and using an intrahuman variability adjustment and distribution for cancer risk derivations. It may sometimes be preferable to use single-value "uncertainty factors," either out of necessity or reflecting science-policy choices (see Chapter 6). Their use would preferably be accompanied by a qualitative description of the associated uncertainty in their application.

  The term *uncertainty factors* can be problematic because it connotes only one aspect of the function of the factors. As the default distributions are developed, a better, more specific label for them would be preferable (for example, *human variability distribution*) to reflect their content more appropriately (for example, accounting for human heterogeneity). This would lessen the opportunity for transferring to the new default distributions the misunderstanding commonly associated with use of "uncertainty factors," as described earlier.
- *Descriptions of sensitive individuals or subpopulations*. The assessment would characterize individuals and subgroups according to whether they have coexposures to key nonchemical stressors, specific polymorphisms influencing metabolism or DNA repair, pre-existing or endogenous disease processes, high background endogenous or exogenous exposures, and other determinants of increased susceptibility.
- *Approaches and resulting assessments that are transparent and understandable by the public and by risk managers*. This may require alternative presentations of the characterization of risk to suit the needs of specific decisions.

Copyright © National Academy of Sciences. All rights reserved.

Science and Decisions: Advancing Risk Assessment
http://www.nap.edu/catalog/12209.html

Risk-Specific Definition of the Reference Dose

This framework facilitates a redefinition of the RfD and RfC in terms of a risk-specific dose and confidence level, as outlined in Box 5-1. Although Box 5-1 focuses on a risk-specific definition of the RfD, the framework developed in this chapter can be used to estimate risk at any dose, not just the RfD; for example, the risk and confidence bounds around the risk could be reported for continuous exposure to an air concentration of 1 part-per-billion. This redefinition will facilitate an understanding of the benefits of lowering exposure in valuation exercises for environmental decision-making.

An RfD defined in that manner can be used as RfDs have always been used in aiding risk-management decisions, but it has additional beneficial features. It presents a dose above which risks may be increased above a standard criterion or de minimis risk and below which risks are considered insignificant or minimal but not necessarily zero. It is analogous to the presentation of cancer risks to risk managers with the understanding that the bright-line risk-specific dose is based on a previously agreed on de minimis or acceptable level of risk inasmuch as zero risk cannot be assumed. However, rather than being an expression of the line between possible harm and safety, the newly defined RfD can be interpreted in terms of population risk. Managers can then weigh alternative options in terms of the percentage of the population that is above or below the de minimis risk-specific dose; this also enables a quantitative estimate of benefits for different risk-management options. An example of this approach is provided for a thyroid disrupting compound by Axelrad et al. (2005).

The de minimis risk for the RfD could depend on the nature of the health outcome (that is, a subtle, precursor effect, a mild effect, or a severe effect) and the subpopulation; for example, the RfD could be based on a 1 in 1,000 risk for a minimally adverse response in a sensitive subpopulation (Hattis et al. 2002).

As is the case for linear cancer end points, multiple risk-specific doses could be provided in the Integrated Risk Information System and in the various risk characterizations that EPA produces to aid environmental decision-making. Different risk-management decisions may call for different acceptable risks, and this redefinition would provide risk managers a means of considering the population risk associated with exposures resulting from specific control strategies. The doses related to different target risks could be distinguished from RfDs and RfCs with names like *risk-specific dose* to avoid confusion. The confidence values associated with these risk-specific doses should be included in any database with the risk targets to ensure that this key information is not lost. Over the years of experience with cancer—a severe effect with a relatively long latent period—an acceptable risk range has been adopted that is used in risk-management decisions. Such experience will accrue for other health end points.

BOX 5-1 A Risk-Specific Reference Dose

For quantal effects, the RfD can be defined to be the dose that corresponds to a particular risk specified to be de minimis (for example, 1 in 100,000) at a defined confidence level (for example, 95%) for the toxicity end point of concern. It can be derived by applying human variability and other adjustment factors (for example, for interspecies differences) represented by distributions rather than default uncertainty factors.

Copyright © National Academy of Sciences. All rights reserved.

Science and Decisions: Advancing Risk Assessment
http://www.nap.edu/catalog/12209.html

Conceptual Models

Approaches to describe dose-response relationships in probabilistic terms depend on how one conceives the underlying biologic processes and how they contribute to an individual's dose-response relationship, the nature of human variability, and the degree to which the processes may be independent of background exposures and processes. This is illustrated in three example prototypical conceptual models:

1. *Nonlinear individual response, low-dose linear population response with background dependence.* As discussed above, low-dose linearity can arise when the dose-response curves for individuals in the population are nonlinear or even have thresholds but the exposure to the chemical in question adds to prevalent background exposures that are contributing to current disease. The dose-response relationship would be determined to a great extent by human variability and background exposure. In Figure 5-4, each individual's dose-response relationship can be characterized by a threshold dose-response function with zero risk up to a particular dose and then sharply increasing risk with increasing dose above it. A collection of the threshold dose-response functions for a number of individuals is displayed on the left side of the figure. The proportion of individuals in the population whose threshold is exceeded by a particular dose is displayed on the right side.

2. *Low-dose nonlinear individual and population response, low-dose response independent of background.* This is the dose-response conceptual model currently in use for noncancer end points. For these dose-response relationships, the fraction of the human population responding drops to inconsequential levels at low doses. At very low doses, the threshold dose for toxicity is not exceeded in individuals, or the risk is infinitesimal. The same is true for the population, with the shape of its dose-response relationship determined by the variability in individuals' thresholds, as illustrated in Figure 5-5.

Clearly, there are many compounds and end points for which available compound-specific data are not sufficient to describe probabilistic dose-response relationships for nonlinear end points adequately. For some chemicals, default distributions may be constructed on the basis of known chemical and physiologic properties for chemicals considered representative for this purpose. Some default adjustment factors could be specific for some types of chemicals. Examples of how default distributions may be derived to support the derivation of risk for this conceptual model are given below; the committee cites these examples not to endorse particular distributions or specific results but to provide an example of a low-dose nonlinear dose-response modeling approach.

FIGURE 5-4 Linear low-dose response in the population dose-response relationship resulting from background xenobiotic and endogenous exposures and variable susceptibility in the population.

Copyright © National Academy of Sciences. All rights reserved.

Science and Decisions: Advancing Risk Assessment
http://www.nap.edu/catalog/12209.html

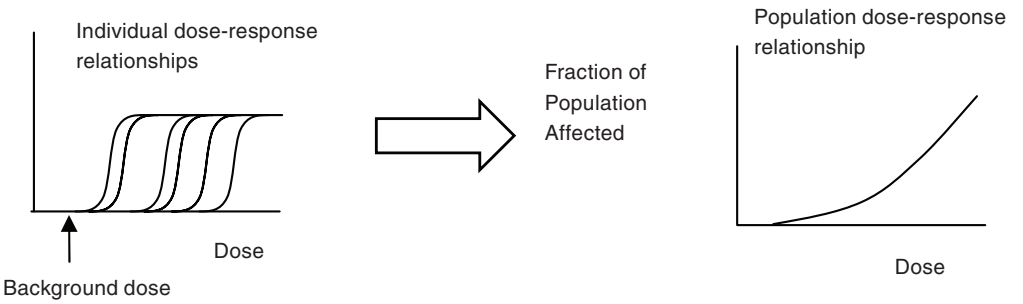

FIGURE 5-5 Nonlinear or threshold low-dose response relationships for individuals and populations.

3. *Low-dose linear individual and population dose-response.* For this conceptual model, both individual risk and population risk have no threshold and are linear at low doses, as illustrated in Figure 5-6. Note that *low-dose linear* means that at low doses "added risk" (above background) increases linearly with increasing dose; it does not mean that the dose-response relationship is linear throughout the dose range between zero dose and high doses. A possible approach for deriving linear cancer dose-response relationships and estimating risk for individuals at different quantiles and for the population is described below for this conceptual model illustrated in Figure 5-6.

To the extent that uncertainty in cross-species and other adjustments can be ascertained, rough quantitative estimates of uncertainty may be provided and incorporated into the characterization of the dose-response relationship. The upper confidence bound on the population dose-response curve in Figure 5-6 depicts the uncertainty in the model fit to data, as well as in the other adjustments.

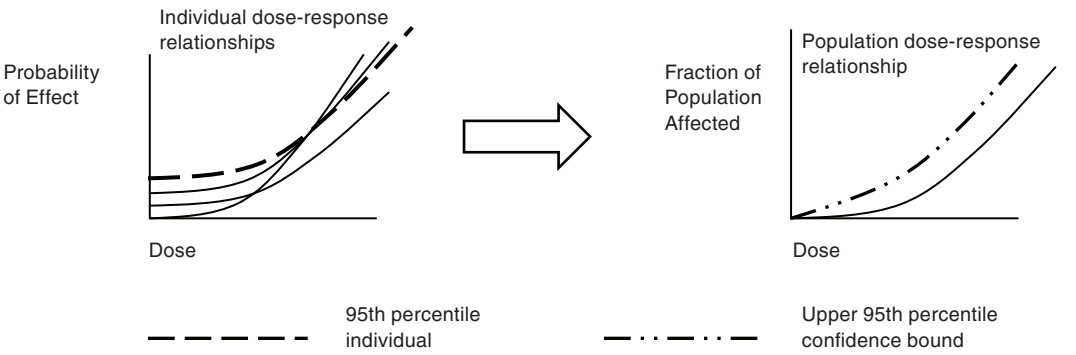

FIGURE 5-6 Linear low-dose response models for individuals and population. Individual dose-response relationships may cross. Thus, individual at the 95th percentile at one dose (dashed line in graph on left) may not be same individual at another dose. From uncertainty estimates for assessment components, upper 95th percentile estimate for population dose-response relationship can be derived (dashed line in graph on right).

Copyright © National Academy of Sciences. All rights reserved.

Science and Decisions: Advancing Risk Assessment
http://www.nap.edu/catalog/12209.html

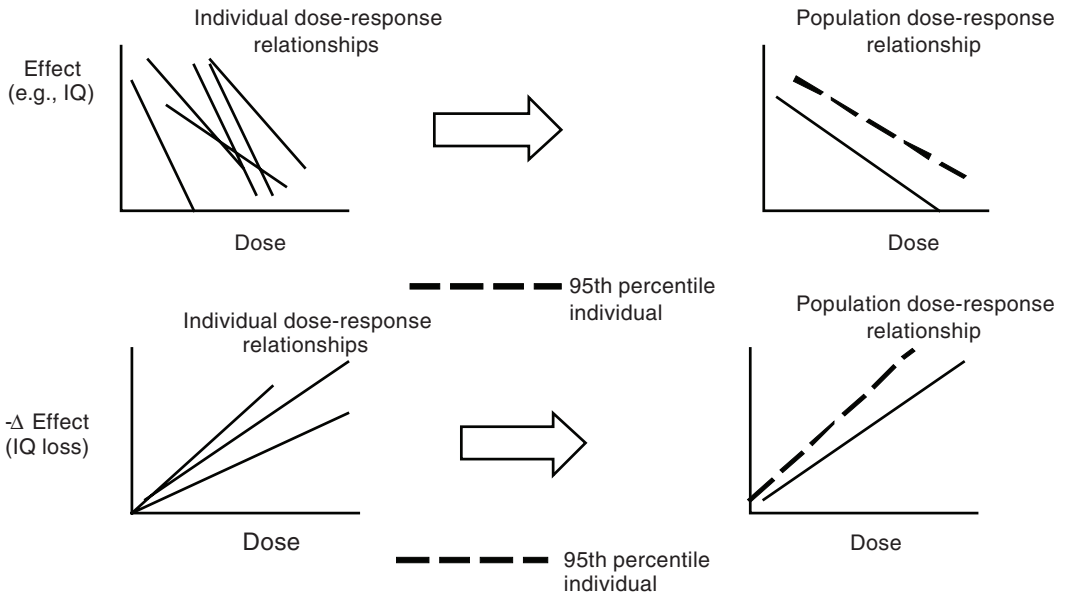

FIGURE 5-7 Dose-response relationships involving a continuous effect variable.

Low-dose linear dose-response relationships can also involve continuous-effect variables, such as decreasing IQ, illustrated in Figure 5-7. As the exposure increases, IQ decreases potentially shifting the entire population distribution in the direction of decreased function, as may occur with methylmercury (Axelrad et al. 2007).

General Approach to Dose-Response Assessment

The general approach, illustrated in Figure 5-8, involves consideration of MOA, background exposures, and possible vulnerable populations in selecting a conceptual model and methods for dose-response analysis.

Data Assembly and End-Point Assessment

The process begins, as is done currently, with review of the peer-reviewed scientific literature to assemble health-effects data for identifying end points of concern. The review emphasizes end points that are of greatest concern to populations exposed through environmental media. Thus, for chemicals with robust datasets, there is little focus on severe effects at high doses other than as indicators, for example, of possible target organs, route specificity, and dose-dependent pharmacokinetics. An exception is the plausible scenario, in which, for example, acute high-dose exposures occur from chemical terrorism or accidental releases.

One important aspect of dataset selection for dose-response estimation is the consideration of target organ (site) concordance between animals and humans. A toxic effect may be preferentially expressed in an animal model in a tissue that is particularly vulnerable because of unique features of metabolism in the tissue, the particular hormonal influences on the tissue, or the rates of aging, damage, and repair in the tissue, and other factors. In some cases, the target organ in a rodent species, such as the forestomach or Zymbal gland, may not have an exact human counterpart. However, the presence of carcinogenic action

Copyright © National Academy of Sciences. All rights reserved.

Science and Decisions: Advancing Risk Assessment
http://www.nap.edu/catalog/12209.html

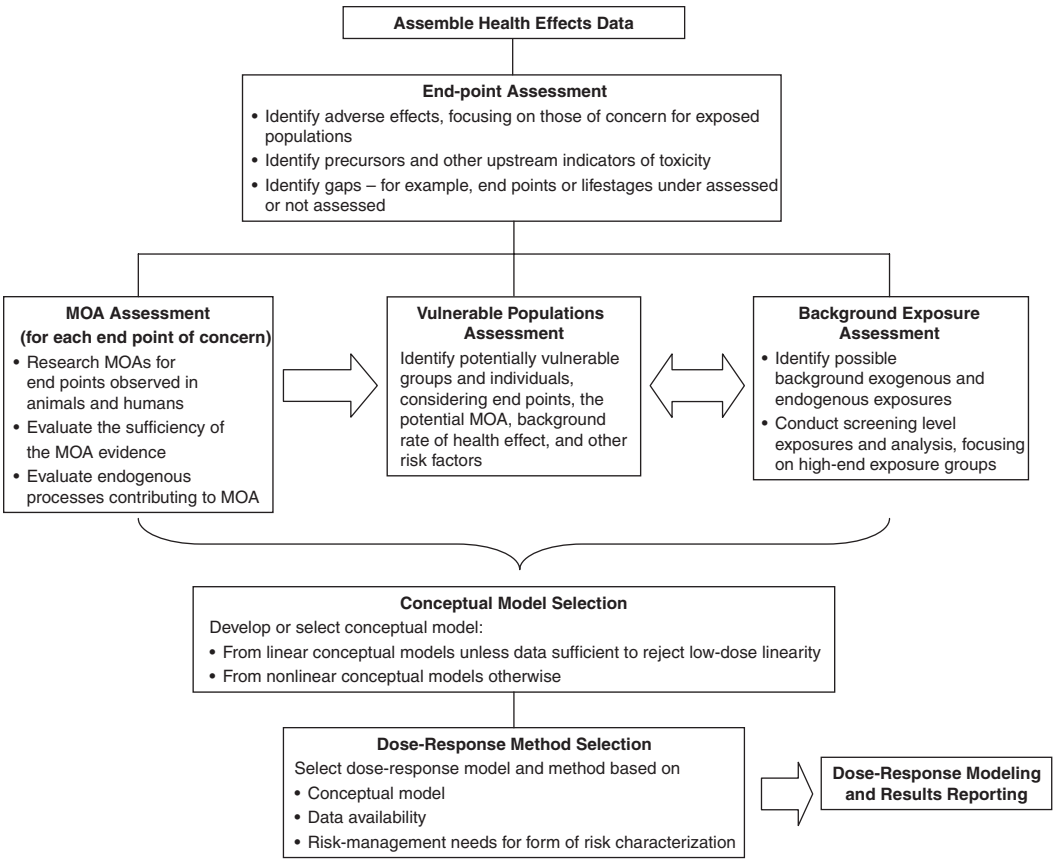

FIGURE 5-8 New unified process for selecting approach and methods for dose-response assessment for cancer and noncancer end points involves evaluation of background exposure and population vulnerability to ascertain potential for linearity in dose-response relationship at low doses and to ascertain vulnerable populations for possible assessment.

in tissues for which there is no correspondence in humans or that may be regulated differently in humans does not mean that the toxicity or tumor finding in animals is irrelevant. That the rodent tissue is sensitive to the toxicant signifies that the toxicant MOAs operate in a mammalian system that has characteristics in common with similar or even not obviously related tissues in humans or human subpopulations. Because epidemiologic studies are often limited in their ability to explore outcomes related to workplace or environmental exposures, it is typically impossible to rule out the relevance of an effect seen in a particular rodent tissue unless there is detailed mechanistic information on why humans would not be affected (IARC 2006). The finding that the high sensitivity of the rat Zymbal gland to benzene tumorigenesis occurs via an MOA (clastogenesis) similar to that which produces benzene-induced bone marrow toxicity and cancer in humans (Angelosanto et al. 1996) is an indication that a tissue that is specific to the rat can still provide important hazard and potency information related to human risk. In general, tissues that are responsive to a toxicant should be considered relevant to human risk assessment unless mechanistic information demonstrates that the processes occurring in the tissues could not occur in humans.

Copyright © National Academy of Sciences. All rights reserved.

Science and Decisions: Advancing Risk Assessment
http://www.nap.edu/catalog/12209.html

Mode-of-Action Assessment

The MOA evaluation explores what is known or hypothesized about the key events after chemical exposures that lead to the toxicity of a compound, including metabolic activation and detoxification, initial interactions with critical cellular targets (for example, covalent binding with protein or DNA, peroxidation of lipids and proteins, DNA methylation, and receptor binding), altered cellular processes (for example, apoptosis, gene expression, and signal transduction), and other types of biochemical perturbation that may involve defense mechanisms or be considered precursor events. Background or endogenous processes that might act in concert with those events would also be considered. Any MOA information that might be helpful in understanding dose-response relationships at both high and low doses would be considered, including dose-dependent nonlinearities in metabolic processes, depletion of cellular defenses, potential to outpace repair processes, induction of enzymes by repeat dosing, additivity and interaction with background disease processes, and additivity of the chemical and its metabolites with other chemical exposures.

The MOA assessment brings mechanistic information to bear on the dose-response assessment. However, the available data will often be too limited to explain how a chemical or its metabolites act to produce an effect. In such cases, default assumptions will apply; below possible defaults are presented in the context of conceptual models. Chapter 6 provides further recommendations and guidance on developing and applying defaults.

Precautionary lessons on the use of MOA data in dose-response assessment are presented by way of the following examples. As the first example, findings of rodent liver cancer have been hypothesized to be of limited or no human relevance for chemicals that are agonists for the peroxisome proliferator activated receptor α (PPARα), a hormone receptor involved in energy homeostasis (Klaunig et al. 2003). Notably findings of rodent liver cancer for di(2-ethylhexyl)phthalate were found by the International Agency for Research on Cancer (IARC 2000) not to be relevant to humans because peroxisome proliferation was demonstrated in mice and rats, but not in human hepatocyte cultures or livers of nonhuman primates exposed to DEHP. However, findings of liver cancer at a higher incidence in PPARα-null than wild-type mice (Ito et al. 2007) call into question this conclusion. Second, MOA assessment has recently been introduced as a way to determine whether a carcinogen has greater sensitivity early in life. Following EPA (2005c), a factor is to be applied when exposure occurs in early life to account for the greater sensitivity during this period, but only for chemicals with established mutagenic MOAs. These guidelines (EPA 2005c) raise the question of what constitutes a mutagenic MOA. It can be difficult to establish how a chemical with some genotoxic activity may induce a mutation (for example, direct vs indirect effect), how to translate findings from one biologic system or age group to another, and how effects are produced when a chemical induces cancer by multiple MOAs, as many carcinogens are likely to do. The practice is inconsistent with the EPA approach to low-dose extrapolation in its cancer risk-assessment guidance: when the MOA is uncertain, the default position is to assume a low-dose linear extrapolation (EPA 2005b, p. 3-21).

The "M" factor described later in this chapter is introduced to modify the dose-response slope at low doses to address the case of multiple MOAs or other aspects that can be different between high and low dose. The MOA assessment would inform the selection of M.

Background and Vulnerability Assessments

A critical aspect of the new approach is the determination that, whether addressing cancer or noncancer end points, dose-response models should fully address both intersubject

Copyright © National Academy of Sciences. All rights reserved.

Science and Decisions: Advancing Risk Assessment
http://www.nap.edu/catalog/12209.html

variability and background disease processes and exposures. How those factors may "linearize" dose-response relationships, which would otherwise be low-dose nonlinear relationships on the basis of MOA, should be considered explicitly. The committee recommends that two systematic reviews be included as components of EPA dose-response assessments. The first is an assessment of background exposures to xenobiotics (for example, in pharmaceuticals, food, and environmental media) and endogenous chemicals that may affect the processes by which the chemical produces toxicity and may result in low-dose linearity. The second is an assessment of human vulnerability that identifies underlying disease processes in the population to which the chemical in question may be adding and that suggests groups of sensitive individuals and their characteristics. Those issues are considered further below in terms of how they may affect the choice of conceptual model used in dose-response analysis.

To facilitate this step of the dose-response assessment process, the committee provides an initial set of diagnostic questions that address whether background considerations are key factors:

- What is known or suspected to be the chemical's MOA?
- What underlying degenerative or disease processes might the toxicant affect or otherwise interact with?
- What are the background incidences and population distributions of these processes?
- Are there identified sensitive populations?
- Have the underlying processes been characterized in humans with markers of susceptibility and precursor effect?
- What known and probable factors can affect the underlying processes and thus potentially modulate adverse health outcomes of exposure to the toxicant?
- What are the levels of human-to-human and age-dependent variability and uncertainty with respect to background degenerative and disease processes, and how do they interact with the toxicant's MOA?
- What environmental contaminants in air, drinking water, food or in consumer products (for example, foods, pharmaceuticals, cosmetics) or endogenous chemicals (for example, natural hormones) are similar to the chemical in question?
- Could they potentially operate by MOAs similar to that of the chemical in question?
- What chemicals might operate by a different MOA but have the potential to affect the same toxic process as the chemical under study?
- How might the endogenous and exogenous background components vary among individuals? Can subgroups with particularly high exposures be identified?
- Is there a potential for people with high background exposures to have health conditions that predispose them to the critical end points or diseases caused by the chemical under study?

Questions, like those above, are essential to ask when conducting chemical risk assessment, whether using the unified framework or current approaches. These questions help identify potential data sources for understanding inter-human variability in response and the extent to which a chemical may pose risks at low doses, and the limits in that understanding. EPA's draft risk assessment for trichloroethylene (TCE) (EPA 2001a; NRC 2006a) took a step in this direction by considering how differences in metabolism, disease, and other factors contribute to human variability in response to TCE, and how other factors may alter its metabolism. EPA's draft dioxin risk assessment considered the impact of background

Copyright © National Academy of Sciences. All rights reserved.

Science and Decisions: Advancing Risk Assessment
http://www.nap.edu/catalog/12209.html

and cumulative exposure to dioxin-like compounds and the potential impact on low-dose response (EPA 2004; NRC 2006b). The unified framework formalizes the incorporation of this type of information into human-health risk assessments, through background and vulnerability assessments and the subsequent selection of a conceptual model for dose-response assessment.

A Pictorial to Aid Vulnerability Assessment

Many factors can affect susceptibility to a chemical, including host genetics, disease status, sex, age, functional reserve, capability of defense mechanisms (for example, glutathione status), capability of repair mechanisms, activity of the immune system, and coexposure to other xenobiotics. Figure 5-9 is an aid to explore how the disease process may be influenced by numerous biochemical processes and risk factors. Someone who is not very vulnerable may have no or few risk factors, whereas someone who is vulnerable may have many or far greater exposure to one or several of them. Figure 5-9 portrays a hypothetical population vulnerability distribution, with the X-axis representing "functional decline," a continuous variable that is an indicator of vulnerability. For example, the indicator of functional decline for asthma could be reduced airway responsiveness. People who have generally lower levels of risk factors and disease precursors will be on the left side of the population distribution in Figure 5-9. Moving to the right will be people who experience a loss of function but are not symptomatic. With further loss of function, as may occur in people who have additional or greater exposure to risk factors, biomarker levels are higher and approach their threshold for symptoms and disease. Stressors that may be innocuous in healthy people may be life-threatening in those who are susceptible. For example, exposure to low concentrations of an infectious agent may cause clinical infection only rarely in the average person, but those whose lung clearance and immune function are compromised may develop pneumonia at a higher frequency and, when afflicted, may have a greater risk of death.

Figure 5-9 illustrates a hypothetical situation in which the population depicted is ex-

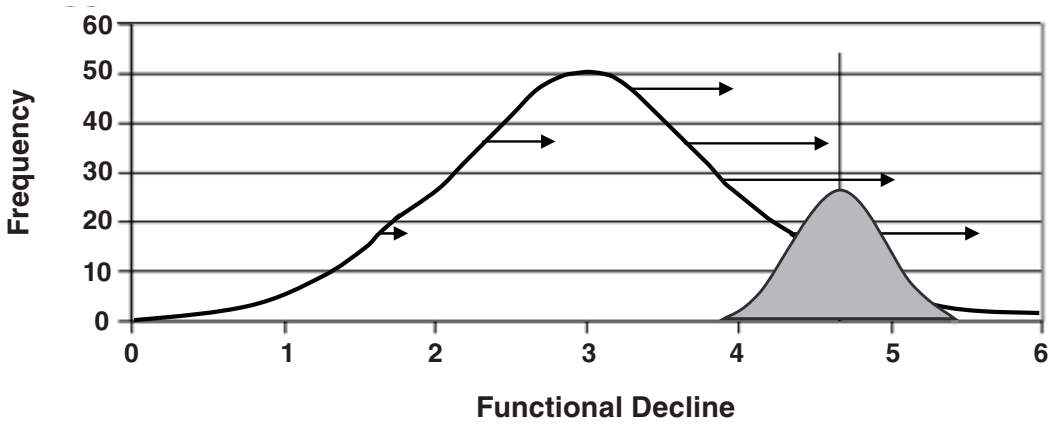

FIGURE 5-9 Population vulnerability distribution. Arrows represent hypothetical response to same toxicant dose for people at given level of functional decline unrelated to any particular toxicant. Vertical line represents presumed threshold between overt adverse and nonadverse effect in median person. Shaded area straddling line represents distribution of thresholds in population.

Copyright © National Academy of Sciences. All rights reserved.

Science and Decisions: Advancing Risk Assessment
http://www.nap.edu/catalog/12209.html

posed to a toxicant. The vertical line represents the theoretical threshold to elicit an adverse clinical effect in the median person. The threshold will not be the same in everyone, so it is represented in the figure as a normal distribution. The arrows represent the magnitude of toxicant effect in response to a given dose in people who are at a given level of functional decline. In this example, people who are more vulnerable are both closer to their threshold and more responsive to a given toxicant dose (represented by the larger arrows). Toxicant exposure will shift the vulnerability distribution to the right and make it more skewed, as indicated by the size of the arrows. Here, as in epidemiology, functional decline or baseline health status might be thought of as an effect modifier of the risk of interest. Sensitivity differs because the more vulnerable, on the one hand, have less functional reserve and cellular defense and, on the other hand, may have a greater number of processes that could contribute to disease (for example, less responsive airways, less pulmonary clearance, poorer immune surveillance, or impaired cardiac function). Low-functioning people can be at greater risk not only because they can be near the threshold but because they can have a greater response per unit dose.

Low doses cause a small shift, and even a very low dose may push a few people over their threshold. If the background level of clinical effect is high (for example, 1% of people have the disease) and there is considerable baseline variability, many people would be expected to be vulnerable to a toxicant-induced increase in the disease. In the case of rare diseases or effects (for example, affecting 1 per 100,000), few people are expected to be just shy of the threshold, and it would take a larger dose of toxicant to produce the same increase in effect as in the high-background case. The diagnostic questions listed above may help the risk assessor to understand the characteristics of the population vulnerability distribution and the potential for low-dose exposures to push some in the population over their threshold.

Selection of a Conceptual Model

Based on the background exposure, MOA, and vulnerability assessments, a decision is made as to the general approach to the dose-response analysis. It involves a selection of conceptual models for individual and population dose-response relationships. To guide this decision, the committee has developed examples of prototypical conceptual models, described earlier and summarized in Figure 5-10.

Consideration of background exposures and processes is critical for the determination of likelihood of low-dose linearity in the population dose-response relationship. Conceptual models 1 and 3 are illustrations of low-dose linearity in population response. The committee recommends that agents be considered as low-dose nonlinear, as in conceptual model 2, only if

- Biologic additivity is not a significant response modifier, for example, there are very low background rates of health end points or damage processes in the population in general, or relevant to the chemical's known or possible MOAs.
- Chemical additivity is not a significant response modifier, that is,
  - the totality of exposure to the toxicant and other agents (exogenous and endogenous) is unlikely to cause the adverse affect, or
  - the toxicant's contribution is so inconsequential that it will not promote the related ongoing toxic processes.

To illustrate the criteria, consider the case of ambient xenon. At high levels, say 70% (mixed with 30% oxygen), xenon is an analgesic and induces a hypnotic effect, and at high

Copyright © National Academy of Sciences. All rights reserved.

Science and Decisions: Advancing Risk Assessment
http://www.nap.edu/catalog/12209.html

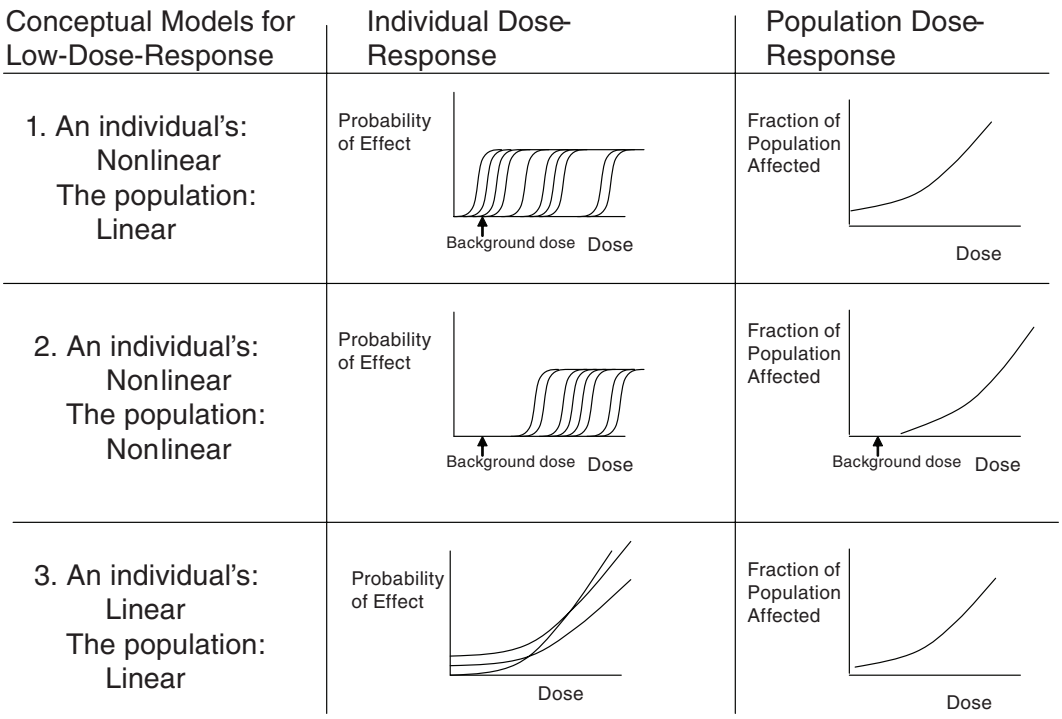

FIGURE 5-10 Examples of conceptual models to describe individual and population dose-response relationships.

levels, xenon displaces oxygen. The MOA for xenon's anesthetic action is unknown but is believed to be electrophysiological in nature, like the other volatile anesthetics. Xenon is ubiquitous in air, at quite a low concentration (0.0000087%). If asked to do a risk assessment for environmental levels of xenon, should a linear or nonlinear approach be applied?

While the MOA is unknown, the number of individuals in the general population with analgesia by xenon relevant MOAs will be restricted to those undergoing surgery, and so the first criterion is met. The totality of exposure to xenon and other volatile anesthetics is not producing anesthesia in the general population. Also, at 0.0000087% xenon's contribution to even those undergoing anesthesia would be inconsequential, as would the degree of oxygen displacement. Thus both criteria point to a threshold approach for the xenon analysis.

Carbon monoxide also impairs blood oxygenation. Its average ambient concentration, expressed as carboxyhemoglobin levels in blood (COHb), is 0.5% COHb. This concentration is less than an order of magnitude below the COHb concentration where effects are observed in human subjects: 2-6% COHb has been associated with increased angina symptoms in those with coronary artery disease. Even in apparently healthy subjects, COHb levels as low as 5% are seen to affect maximal exercise time and the maximal exercise level. Furthermore, concentrations of carbon monoxide in air can fluctuate diurnally, geographically, and by activity (for example, driving). Thus in evaluating the risk of carbon monoxide exposure, both of the above criteria indicate a linear approach should be considered: coronary heart disease is common and increased carbon monoxide exposures will likely contribute to ongoing toxic processes.

The recommendation to consider background exposure and vulnerability in deciding

Copyright © National Academy of Sciences. All rights reserved.

Science and Decisions: Advancing Risk Assessment
http://www.nap.edu/catalog/12209.html

between linear and low-dose nonlinear approaches applies even to agents that, when tested in isolation in rodent models, appear to have a threshold and whose MOAs (in the absence of consideration of background and human heterogeneity) would otherwise suggest a threshold. Approaches and guidelines for conducting vulnerability and background assessments will be needed, as will guidelines for conducting the assessments and selecting conceptual models.

Selection of Method for Dose-Response Analysis

The approach to the analysis depends on the conceptual model, the data available for the analysis, and risk-management needs. If, for example, data are sparse and available only from animal studies and low-dose linearity is ruled out, the analysis may proceed by using default distributions for adjustment factors and using methods like those described in the next section. If there is a relatively high endogenous or exogenous background exposure to the same and related chemicals or vulnerability can be substantial and highly variable (perhaps in particularly sensitive subgroups), the analysis may proceed by a linear default or incorporate distributional information specific to the particular chemical or circumstance being analyzed.

The following section suggests approaches to dose-response analyses for a variety of toxic mechanisms and interactions with background processes and exposures. The general assumption in working through the examples provided is that variability distributions are unimodal: people who are at an extreme for a particular parameter are not numerous enough to constitute a subpopulation that should be analyzed separately. However, for any given parameter (for example, respiratory function, immunoglobulin E status, blood pressure, xenobiotic-metabolizing capacity, or DNA repair), a multimodal distribution may exist and be influential enough to create a multimodal distribution of risk at a given dose.

Unique subpopulations can be addressed as special cases within the framework. Figure 5-11 depicts such a case, showing that the dose-response relationship for sensitive people has very little overlap with that for the typical person. If the sensitive people constitute a distinct group either because of their numbers or because of identifiable characteristics—such as ethnicity, genetic polymorphism, functional or health status, or disease—they should be considered for separate treatment in the overall risk assessment. An example of a generally susceptible well-defined group is asthmatics, with respect to their response to irritant gases emitted from rocket engines (NRC 1998a). Analysis of dose-response functions of asthmatic subjects indicated sensitivity to hydrochloric acid potentially 3 times greater, to nitrogen dioxide 10 times greater, and to nitric acid 20 times greater than healthy individuals, respectively. The committee reviewing the data considered that a multimodal distribution that includes the variance and distributional form within each mode was needed for full characterization of the range of sensitivity to those irritants. Issues of threshold and background additivity can be analyzed separately for each mode to determine whether low-dose linearity assumptions are appropriate for one or more subpopulations. While consideration of susceptible subpopulations has been included in a number of environmental risk assessments (for example, NRC 2000 [copper and Wilson's disease heterozygotes]; EPA 2001b [methylmercury effects on developing children]), the level of consideration and incorporation in EPA assessments could be much improved. The conceptual framework and committee recommendations in this chapter support qualitative and quantitative improvements.

Copyright © National Academy of Sciences. All rights reserved.

Science and Decisions: Advancing Risk Assessment
http://www.nap.edu/catalog/12209.html

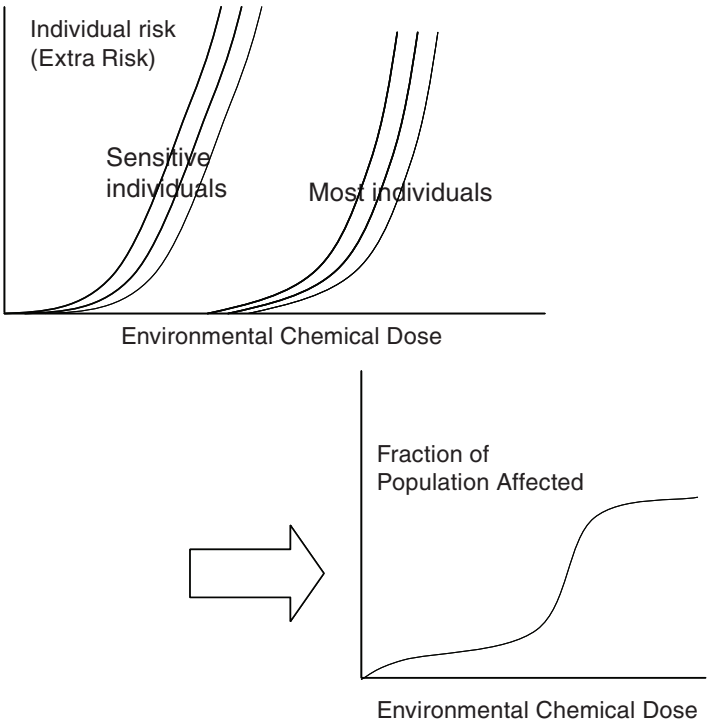

FIGURE 5-11 Widely differing sensitivity can create a bimodal distribution of risk.

CASE STUDIES AND POSSIBLE MODELING APPROACHES

This section provides case studies and possible methods for dose-response analysis for the three example conceptual models, as outlined in Figure 5-12. Methods take into account the nature of the data available. Some methods are "bottom up" in that the dose-response relationship is constructed from components. An example is given for how human variability in asthmatic response might be inferred from gene polymorphisms and might lead to a description of the population dose-response relationship for asthma. Other methods are "top down" in that the dose-response relationship at low doses is derived by fitting exposure-response models to observations from epidemiologic or animal studies.

Conceptual Model 1: Low-Dose Linear Dose-Response Relationship Due to Heterogeneous Individual Thresholds and High Background

Particulate-Matter Case Study

Fine PM ($PM_{2.5}$) belongs to a family of pollutants (including ozone) with noncancer end points for which the evidence points to a linear or other nonthreshold population response at low doses. For those agents, exposed individuals have different thresholds, and full characterization of the distribution of thresholds in the population (in this case based on epidemiologic evidence) is informative for a population concentration-response function. Numerous factors contribute to the distribution of the thresholds, as explained later.

Copyright © National Academy of Sciences. All rights reserved.

Science and Decisions: Advancing Risk Assessment
http://www.nap.edu/catalog/12209.html

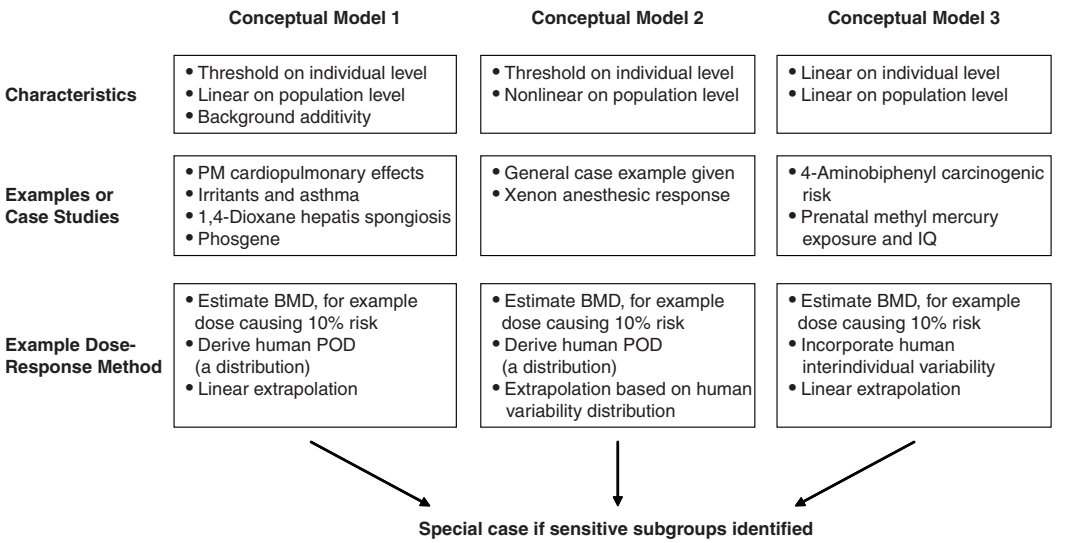

FIGURE 5-12 Three example conceptual models lead to different descriptions of dose-response relationships at individual or population levels. These are illustrated in the case studies. For each conceptual model, there may be a sensitive subgroup that should be addressed with separate dose-response analysis.

Furthermore, $PM_{2.5}$ is an example of pollutants that have numerous sources of exposure, so any analysis of a given source of $PM_{2.5}$ takes place against a background that may already be above a threshold for numerous people.

This case illustrates two dose-response issues that are of particular interest to the committee's framework for dose-response assessment:

- How concentration-response functions are developed throughout the range of observed exposures, taking into account potential nonlinearities and population thresholds.
- How human heterogeneity in response has been quantified and formally addressed both to understand sensitive subpopulations and to determine the distribution of individual thresholds to understand low-dose effects better.

How concentration-response functions are determined outside the range of observed exposures is not addressed. The available epidemiologic evidence for $PM_{2.5}$ analyses has involved fairly low-level exposures, and extrapolation below the level of observation to any great degree is less important than for compounds for which evidence is derived from animal bioassays or occupational (high dose) epidemiology.

The $PM_{2.5}$ dose-response assessment entails the construction, from epidemiologic observations, of a concentration-response function spanning all observed levels of exposure. Such a function could be used to determine directly the proportion of people whose thresholds were exceeded by a given concentration (as described above), if concentration-response functions were developed all the way down to the lowest observed exposure (ideally, approaching nonanthropogenic background). However, in a benefit-cost analysis framework, the question of the slope of the concentration-response curve near nonanthropogenic background is irrelevant because any feasible control strategies involve incremental exposure reductions and some residual exposure. For the $PM_{2.5}$ case, an important outcome of the assessment for

Copyright © National Academy of Sciences. All rights reserved.

Science and Decisions: Advancing Risk Assessment
http://www.nap.edu/catalog/12209.html

risk management is the difference in the proportions of people adversely affected between a precontrol and a postcontrol scenario. Thus, the analysis has focused on risks in regions of the dose-response curve in which control options are relevant.

Some investigators have used statistical techniques to investigate whether any nonlinearities (including population thresholds) were present in $PM_{2.5}$ concentration-response functions in the range of observed data. For time-series studies looking at mortality and morbidity end points, the statistical methods used have included generalized additive models (Schwartz et al. 2002) and penalized regression splines (Samoli et al. 2005). Other studies have evaluated the questions of thresholds and nonlinearities explicitly by fitting piecewise linear concentration-response functions with defined knot points and then using model averaging based on the posterior probabilities of the various candidate models (Schwartz et al. 2008). Regardless of the approach, any of these techniques allow the explicit consideration of nonlinearities in concentration-response functions, including the possibility of population thresholds. However, these approaches are clearly applicable only to epidemiologic evidence, in which there are observations at a sufficient number of magnitudes of exposure to infer the shape of the concentration-response function empirically rather than on the basis of prior hypotheses about functional form. It is also most relevant for population rather than occupational epidemiology, so it will be valuable for only a small number of compounds (those to which exposure is ubiquitous and which pose relatively high population risks).

One crucial question is whether those statistical methods have demonstrated population thresholds for $PM_{2.5}$ or substantial departures from linearity. Another is whether the data would ever be rich enough to discriminate between a model with a threshold and a model without a threshold. Most studies that have used the methods (Schwartz and Zanobetti 2000; Daniels et al. 2000; Schwartz et al. 2002; Dominici et al. 2003; Samoli et al. 2005) have concluded that the functions are effectively linear throughout the range of observed concentrations, which, in the case of many time-series studies, approaches zero. Thus, in spite of the use of statistical models that could detect population thresholds, or at least low-dose nonlinearity, no thresholds appeared to be present in the range of observed concentrations. That finding has been attributed (Schwartz et al. 2002) to the fact that there is a wide distribution of individual thresholds and, in the case of cardiopulmonary mortality (a background disease process with which $PM_{2.5}$ exposures are associated), numerous genetic, environmental, disease-state, and behavioral risk factors each contribute to the distribution of the thresholds.

The extent of the distribution of individual thresholds was quantified by one study of PK and PD factors that influence heterogeneity in response to $PM_{2.5}$ (Hattis et al. 2001). The study assumed lognormality to describe the distribution for individual thresholds. The study concluded that the most susceptible (99.9th percentile) people would respond at doses only 0.2-0.7% of those needed to exhibit responses in people of median susceptibility. An extension of this analysis found results for subpopulations that were consistent with lognormal distributions for a very small number of cut points (Hattis 2008), suggesting the general population responses may be consistent with a mixture of lognormal distributions. Given that the analyses did not include all important aspects of coexposures and disease states that might influence vulnerability, the true heterogeneity could be greater. That provides good physiologic plausibility of low-dose linearity on a population basis, given ubiquitous exposures that imply that a substantial number of people will be found to be at least as sensitive as the 99.9th percentile individual.

Human heterogeneity in response has also been evaluated epidemiologically through the examination of effect modifiers to identify sensitive subpopulations. For example, multiple studies have found that the relative risk of cardiovascular end points (ranging from markers

Copyright © National Academy of Sciences. All rights reserved.

Science and Decisions: Advancing Risk Assessment
http://www.nap.edu/catalog/12209.html

of systemic inflammation to hospitalization to death) was increased in people with diabetes, hypertension, or conduction disorders of the heart (Zanobetti et al. 2000; Dubowsky et al. 2006; Peel et al. 2007). In principle, such pooled evidence from multiple studies could allow a calculation of the risk of an effect of a defined dose in subpopulations with and without specific conditions. Instead of attempting to define risk-specific doses for a pooled population that includes a wide range of sensitivities, a stratified analysis could be performed of the range of thresholds possible in the population on the basis of what is known about unique and definable subgroups.

There are some aspects of $PM_{2.5}$ and other criteria pollutants that are not generalizable to other pollutants, but this case example illustrates the greater role that epidemiology could play in unified toxicity assessments. Opportunities to develop concentration-response functions for noncancer end points should be exploited by using statistical techniques to draw empirical inferences about the shape of the concentration-response function in the range of observed data, taking account of sensitive subpopulations. This case also serves as a reminder that EPA is already developing quantitative risk estimates for a few noncancer stressors that go beyond the threshold concept and has been doing so for some time.

Asthma Case Study

The $PM_{2.5}$ case provided an example of how "top-down" methods can be used to characterize the population distribution of vulnerability. "Bottom-up" approaches may also be informative, as described by this example. These approaches entail characterization of background processes of function loss, damage, disease, and concomitant exposures that will enable a description of the population distribution of vulnerability. That, in turn, can be used in assessing interindividual variability in toxicodynamic response at low doses and can inform the shape of the dose-response relationship at low doses. A case study of asthma is used to explore the concept. Here evidence from markers of disease susceptibility combined with analyses of genotypic differences in vulnerability and relatively high background asthma incidence are considered to evaluate the potential for asthmagenic chemicals to have linear-dose-response relationships at low doses.

Host markers of susceptibility to asthma have been developed and can be used to construct a vulnerability distribution. Asthma occurs in people who are hyperresponsive to allergens and irritants and are thus at the high end of the population distribution of airway responsiveness. The methacholine-challenge test is one of several probes used to screen populations for airway reactivity and used in the diagnosis of asthma. Methacholine is a cholinergic bronchoconstrictor in both normal and hyperreactive airways; there is a continuous distribution of airway reactivity as defined by the challenge dose required to decrease FEV_1 by a given percentage. FEV_1 is the volume of air that can be forced out of the lungs in 1 s after a person takes a deep breath. The PC_{20} is the provoking concentration of methacholine required to decrease FEV_1 by 20%. Among healthy, nonasthmatic people, this measure is distributed so that the majority have low reactivity (high PC_{20}) and a subset have high to very high reactivity. The PC_{20} of 8 mg/L has been used as a cut point to indicate airway hyperreactivity; a person with a PC_{20} below this value is considered to be hyperresponsive and is likely to be either asthmatic or vulnerable to becoming asthmatic. Those with reactive airways appear to be at increased risk for xenobiotic triggering of symptoms and the onset of clinically diagnosed asthma, as indicated in prospective studies that contrast "normoresponders" with asymptomatic "hyperresponders" (Laprise and Boulet 1997; Boutet et al. 2007). The hyperresponders tended to develop more asthmatic symptoms and have decreasing PC_{20}.

Copyright © National Academy of Sciences. All rights reserved.

Science and Decisions: Advancing Risk Assessment
http://www.nap.edu/catalog/12209.html

Boutet et al. (2007) evaluated the distribution of PC_{20} values in a population of 428 healthy vocational students in the province of Quebec, Canada. Figure 5-13 is constructed from the data presented in that study. Asymptomatic hyperresponsiveness (PC_{20} less than 8 mg/mL) was observed in 8.5% of the subjects. The increase in respiratory symptoms over a 3-year observation period differed dramatically in this population. Those most at risk had the highest baseline response to methacholine (PC_{20} less than 4 mg/mL); these high responders had a relative risk of symptoms of over 30 compared with baseline normal responders (PC_{20} over 32 mg/mL). The increase in symptoms in this population was apparently not related to workplace exposure and so may reflect a generalized trend toward the asthmatic phenotype in otherwise healthy people who are asymptomatic hyperresponders in the initial screening. This finding is reinforced by a similar earlier occupational study of animal workers and bakers (de Meer et al. 2003).

The findings indicate how an underlying disease factor, such as airway hyperresponsiveness, can influence the onset of new disease (in this case asthma) in the population. The more people are in the asymptomatic but vulnerable range, the more likely it is that new cases of disease will occur. Different populations may have different background distributions of predisposing risk factors, as shown in an analysis of PC_{20} data by Hattis (2008).

The background rate of airway hyperresponsiveness may be used to assess the number of people at risk for developing asthma symptoms in response to even low doses of a new insulting agent. If the background rate of hyperresponsiveness is low, the number of people near the threshold for symptoms may also be low, and the low-dose incremental effects of the toxicant may have a linear dose-response relationship but with a shallow slope. If many people are vulnerable, the slope at a low dose may be steeper, with a greater incremental effect increase per unit of exposure. Thus, variability in this precursor characteristic, airway hyperresponsiveness, may be a key input into a distributional analysis of the effects of ozone or other toxicants on asthma risk. It will be a challenge to toxicology and epidemiology to generate data that can inform understanding of the interaction of toxicants with predisposing disease factors in vulnerable populations. A simplistic approach to these relationships for asthma follows.

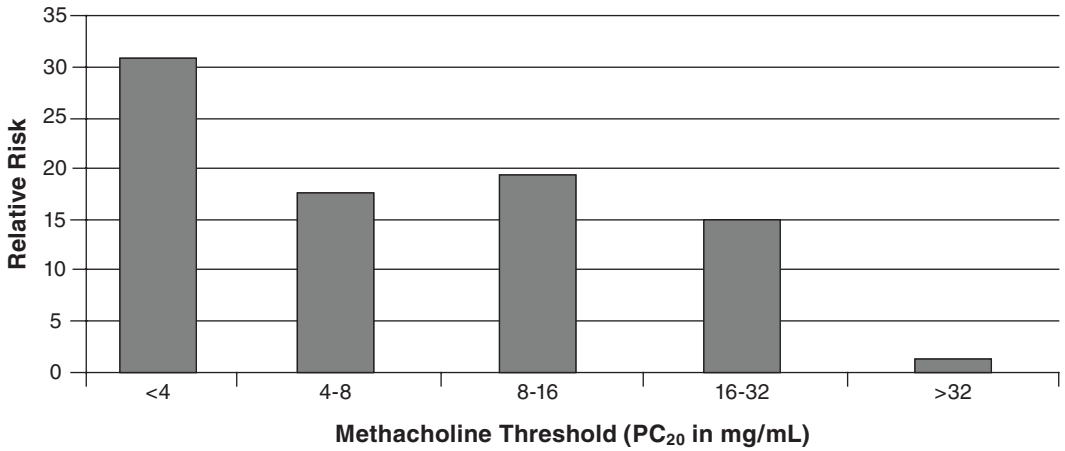

FIGURE 5-13 Baseline airway reactivity as vulnerability factor for allergen-induced respiratory effects expressed as relative risk. Source: Data from Boutet et al. 2007. Reprinted with permission; copyright 2007, *Thorax*.

Copyright © National Academy of Sciences. All rights reserved.

Science and Decisions: Advancing Risk Assessment
http://www.nap.edu/catalog/12209.html

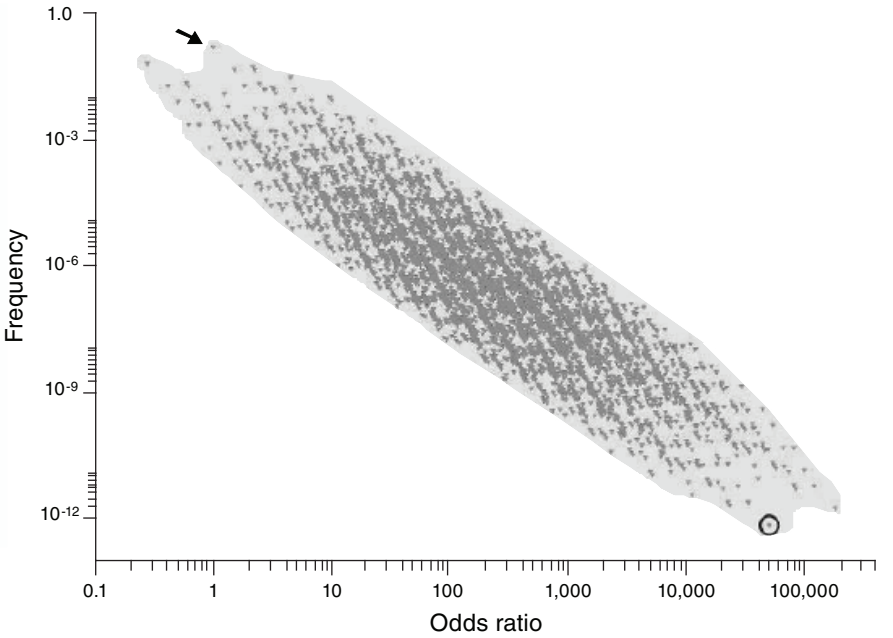

FIGURE 5-14 Effect of asthma-related gene polymorphisms on human vulnerability to asthma. Odds ratios and frequencies were calculated with assumption of 16 gene variants, each point representing a unique combination. Referent genotype (arrow) has odds ratio of 1, and profile composed of all variants (circle) is at other extreme. Source: Demchuk et al. 2007.

Progress has been made in identifying specific genetic factors that predispose to asthma. A recent publication breaks the factors into three broad categories: immune and inflammatory (12 genes), atopic (three genes), and metabolic (one gene) (Demchuk et al. 2007). By accounting for the population frequency of polymorphisms that affect gene expression or protein function and for the odds ratio associated with each polymorphism in terms of asthma risk, the analysis provided a population distribution of vulnerability to asthma, as shown in Figure 5-14. If some people had all the higher-risk polymorphisms (circle) and the sensitivity-enhancing effects acted multiplicatively when combined, these people would have a roughly 50,000-fold increase in risk of developing asthma compared with all wild-type people (arrow). Kramer et al. (2006) propose ways of identifying key candidate genes to better describe genetic susceptibility on PM induced asthma and how research might better support the regulatory standard-setting for PM. Modeling exercises can explore toxicant interaction with the polymorphic pathways to see how exposure in conjunction with host variability may combine to create a distribution of risk of asthma. In the absence of such an understanding, it would be reasonable to assume that chemicals that induce or exacerbate asthma do not have threshold dose-response relationships at the population level and that low-dose linearity prevails.

1,4-Dioxane in Animals Case Study

When epidemiologic data are lacking, diagnostic questions regarding vulnerability and background exposures may be difficult to answer. The background rate of toxicity in

Copyright © National Academy of Sciences. All rights reserved.

Science and Decisions: Advancing Risk Assessment
http://www.nap.edu/catalog/12209.html

unexposed animals and the shape of the dose-response relationship may indicate whether background or endogenous processes will be important in evaluating the potential for low-dose linearity. Variability is expected to be much greater in the human population than in tester strains bred for use in the laboratory and exposed under controlled conditions, so it is important to reflect on potential human processes in reaching overall conclusions. However, animal studies can be more thorough in evaluating age-related and spontaneous toxicity in the control group than is typically possible in unexposed or reference human populations. Therefore, animal toxicity studies may provide important insights into the potential for low-dose linearity.

A case in point is 1,4-dioxane. This solvent produces histopathologic changes in the liver's Ito cells termed hepatic spongiosis—an inflammatory lesion of the sinusoidal and endothelial cells that can be progressive and is believed to be involved in the response to nitrosamines and other hepatocarcinogens in rodents (Karbe and Kerlin 2002; Bannasch 2003). This end point is sensitive to 1,4-dioxane exposure (Yamazaki et al. 1994) and is an example of a noncancer end point. However, evidence of its involvement as a precursor lesion in hepatocarcinogenesis could lead to its evaluation with a different analytic framework (for example, conceptual model 3). As shown below, control males have a high incidence (24%), whereas this lesion was not detected in the control and lowest-dose females. The sex-specific differences in background incidence of and sensitivity to liver disease mirror the pattern of hepatocarcinogenesis in rats and humans, with males more commonly affected than females (West et al. 2006).

As seen in Figure 5-15, the high background rate of the toxic end point in males is associated with a steeper dose-response curve at low dose in males than in females; this is consistent with the shape of the dose-response curve expected on the basis of the background rate of response.

The potential for background processes to affect the shape of the dose-response curve for specific toxicants as observed in animal studies should be considered in building PD variability distributions in humans and in evaluating the possibility of low-dose linearity. In the case of the hepatic effect caused by 1,4-dioxane, prefibrotic and precirrhotic findings

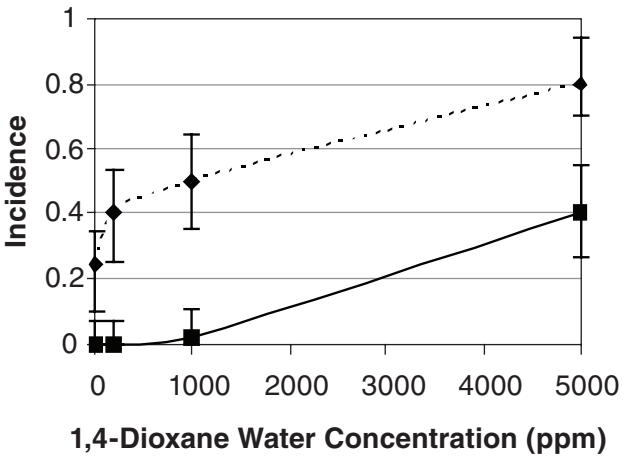

FIGURE 5-15 Dose-response relationship for liver spongiosis in 1,4-dioxane-exposed rats. Bars indicate the 95% confidence intervals. Source: Adapted from Yamazaki et al. 1994.

Copyright © National Academy of Sciences. All rights reserved.

Science and Decisions: Advancing Risk Assessment
http://www.nap.edu/catalog/12209.html

in the human population would be helpful in weighing the relevance of findings on animal vulnerability to that likely to occur in people. Diagnostic methods that can detect subtle liver damage in humans, such as ultrasonography and liver-function tests, may help in exploring background vulnerability to hepatotoxicants if developed further and applied to populations of healthy people (Hsiao et al. 2004; Maroni and Fanetti 2006). Existing underlying conditions and their causes could be considered in the context of potential mechanisms of 1,4-dioxane toxicity to evaluate whether the dose-response relationship should be treated as linear or nonlinear at low doses.

Default Modeling Approach for Conceptual Model 1: Linear Extrapolation for Phosgene

As described above, small chemical exposures in the presence of existing disease processes and other endogenous and exogenous exposures can have linear dose-response relationships at low doses. Thus, a simple methodologic default to address conceptual model 1 compounds is linear extrapolation from the POD, such as a benchmark dose, down to low doses. Greater information on MOA and chemical interactions with background disease processes and similarly acting chemicals may allow different low-dose extrapolations. For example, the slope of the line at the POD or another particular dose could be adjusted, as described below for conceptual model 3.

Linear low-dose extrapolation for a noncancer end point is illustrated with the case example of phosgene. This reactive respiratory toxicant damages the airways at high doses, and dose-response studies in rats exposed for 12 weeks report effects of inflammation and fibrosis of the bronchiolar region (Kodavanti et al. 1997; EPA 2005d). The BMD_{10} for this phosgene effect in rats is 170 $\mu g/m^3$ as a human equivalent concentration (HEC). The lower 95% confidence bound—the $BMDL_{10}$—is 30 $\mu g/m^3$. To this an adjustment is made because the study is subchronic rather chronic, and chronic exposure is of interest in calculating an alternative RfD.

In considering how this risk may be manifested in human populations, the background incidence of asthma—about 6% in children (CDC 2007)—is relevant. Asthmatics experience inflammation, fibrosis, and airway remodeling in response to environmental allergens and irritants and so constitute a large population potentially vulnerable to phosgene. In addition there are numerous medical conditions (for example, infection, environmental exposures, and pharmaceuticals) that lead to the lung inflammation and fibrosis that would potentially be worsened by phosgene exposure. Thus, there is a potential for background additivity that is consistent with conceptual model 1 and linear extrapolation to low dose. Further analysis of cell types and disease processes involved in phosgene toxicity and the other medical conditions may lead one to discover otherwise, but absent more definitive information indicating implausibility, background additivity would be assumed.

Box 5-2 shows that a linear extrapolation from the BMD derived by EPA would yield a risk-specific dose (median estimate) of 0.0085 $\mu g/m^3$ phosgene exposure. Theoretically, exposure at this dose is predicted to contribute to inflammation and fibrosis in 1 in 10^5 of exposed individuals. The phosgene RfC of 0.3 $\mu g/m^3$, set by EPA with a 100-fold cumulative uncertainty factor, corresponds to a theoretical risk that 1 in 3,000 (median estimate) individuals could be affected, on the basis of linear extrapolation. Implicit in the extrapolation are the assumptions that a 10-fold reduction in exposure will result in a 10-fold reduction in risk and that the $BMDL_{10}$ in terms of the HEC is the human 10% effect dose. This approach could be refined to explore the variability between individuals that is possible because of pharmacokinetics, the incidence and distribution of relevant respiratory health conditions, and many other factors, and to explore issues regarding species dose-effect concordance for

Copyright © National Academy of Sciences. All rights reserved.

Science and Decisions: Advancing Risk Assessment
http://www.nap.edu/catalog/12209.html

BOX 5-2 Conceptual Model 1:
Default Linear Low-Dose Extrapolation for Phosgene

1. Assume uncertainty in all parameters can be characterized by a lognormal distribution, with standard deviation represented by σ.
2. BMD_{10} (human equivalent concentration) = 170 $\mu g/m^3$, with 95%-tile lower bound 30 $\mu g/m^3$ variability in animal BMD, with a difference between lower 95% bound and median of 5.7-fold (because 5.7=170/30):
   $\sigma_{Animal\ BMD} = \log(5.7)/1.645 = 0.46$
   (Division by the 95% confidence bound is 1.645 standard deviations from the median in the standard normal distribution.)
3. The human equivalent concentration accounts for cross-species differerence in pharmacokinetics but not pharmacodynamics.
   Assume, as in Hattis et al. 2002, that $\sigma_{logA\rightarrow H} = 0.42$
4. Median human POD:
   Adjust for subchronic to chronic study length, as in Hattis et al. 2002, by a factor of 2:
   170 $\mu g/m^3 \div 2 = 85\ \mu g/m^3$
   Assume the uncertainty ($\sigma^2_{logSC\rightarrow C}$) in the adjustment, as in Hattis et al. 2002:
   $\sigma_{logSC\rightarrow C} = \log[2.17] = 0.34$
5. Uncertainty in the human POD ($\sigma_{logHuman\ POD}$):
   $\sigma^2_{logHuman\ POD} = \sigma^2_{logAnimal\ BMD} + \sigma^2_{logA\rightarrow H} + \sigma^2_{logSC\rightarrow C}$
   $\sigma^2_{Human\ POD} = 0.46^2 + 0.42^2 + 0.34^2 = 0.71^2$
6. Lower 95% confidence bound on Human POD =
   (median human POD)/$10^{[(1.645)(\sigma logHuman\ POD)]}$ = $85/10^{[(1.645)(0.71)]} = 85/14.7 = 5.8\ \mu g/m^3$
7. Linear extrapolation to risk-specific dose - inflammation of 1 in 10^5 people would be affected:
   risk-specific dose = $10^{-5} \times (85/0.1) = 0.0085\ \mu g/m^3$, with lower bound $0.00058\ \mu g/m^3$
8. Estimate risk at different doses: for example, at 0.01 $\mu g/m^3$, three people in 10^5 (median estimate) would be affected.

phosgene. Here, as for conceptual model 3, an important issue is whether dose effectiveness is the same at high doses and low doses. Extrapolation methods for addressing that are discussed in the section below on the mathematical framework for conceptual model 3.

Conceptual Model 2: Low-Dose Nonlinear Dose-Response in Individuals and the Population, Low-Dose Response Independent of Background

The approach would be applied when there is sufficient evidence to reject the possibility of low-dose linearity on the basis of vulnerability and background assessments. As discussed above, the committee encourages the agency to conduct the necessary research and develop appropriate methods and practices for using probabilistic methods for low-dose nonlinear end points. To illustrate the approach, an example methodology using distributions for making calculations is laid out here and sample calculations are applied for a general case. The committee acknowledges that work is needed to further develop the underlying distributions and that methods are needed to support their use in a regulatory context.

Deriving a Reference Dose with Probabilistic Methods

Published methods and examples describing noncancer risk probabilistically (Gaylor et al. 1999; Evans et al. 2001; Hattis et al. 2002; Axelrad et al. 2005; Hattis and Lynch 2007;

Copyright © National Academy of Sciences. All rights reserved.

Science and Decisions: Advancing Risk Assessment
http://www.nap.edu/catalog/12209.html

Woodruff et al. 2007) illustrate a general approach or elements of it that can be used for this conceptual model. They can lead to an RfD based on a de minimis risk target, such as a specified fraction of the population exceeding a threshold, and the uncertainty in that estimate (for example, less than 1 in 100,000 people with some threshold response with a 95% confidence interval).

The general approach is to use distributions for the adjustments to the POD to derive a human-based POD and then to extrapolate from the human POD to lower doses on the basis of assumptions about how humans differ in susceptibility. Figure 5-16 shows how the adjustments from the animal to the human POD could be made. They are depicted here as distributions for the subchronic-to-chronic adjustment (if animal study was of less than chronic duration), database deficiencies, and animal-to-human adjustment, encompassing PK and PD across-species variability. As illustrated in Figure 5-16, the adjustment distributions can be convolved by using statistical or numerical approaches to form an overall adjustment and uncertainty distribution. Quantitatively accounting for the uncertainty in the adjustments enables a quantitative expression of the uncertainty in the overall adjustment. The adjustment distribution is applied to the animal POD to derive a distribution for the human POD. The extrapolation from the POD down the dose-response curve is driven by interhuman variability, broken out in Figure 5-16 into PK and PD elements. The application of adjustment and uncertainty distributions representing each of these elements effectively converts the animal POD (for example, the BMDL or the ED_{50}, the effective dose estimated to affect 50% of subjects) to a probabilistic dose-response relationship for the human population with confidence bounds based on the adjustment distributions.

It is possible in principle to derive the RfD on the basis of some upper percentile value selected from each of the distributions. That would yield a single estimate, similar to the current approach. The preferred method is to incorporate the full distributional information on each component factor by using probabilistic approaches, such as a Monte Carlo approach or a simple analytic approach (for example, when adjustments can be described by lognormal distributions). In that case, the RfD could be selected as a confidence point on the probability distribution for the fraction of the population with a defined risk. Alternatively, the population risk posed by a given dose could be described with a probability distribution reflecting the uncertainty in the estimate.

The approach relies on distributions for the adjustment factors. Researchers developing the method have defined distributions of each of the factors from empirical databases, as briefly summarized below. These distributions are provided to show how they might be derived, not as an endorsement of any specific distribution or their use by EPA. The distributions that lead to the adjustment of the animal POD to the human POD are described first, and then those used to extrapolate from the human POD to lower doses.

Distributions to Adjust Animal POD to Human POD

- *Subchronic-to-chronic factor*. Subchronic and chronic NOAELs from a database of 61 chemicals were compared and statistically analyzed (Weil and McCollister 1963; Nessel et al. 1995; Baird et al. 1996). A lognormal distribution was fitted to the data, which had a geometric mean of 2.01 (that is, the subchronic NOAEL was generally twice the chronic NOAEL) and a geometric standard deviation of 2.17 (Hattis et al. 2002). The standard 10-fold adjustment factor for subchronic-to-chronic extrapolation was about at the 98th

Copyright © National Academy of Sciences. All rights reserved.

Science and Decisions: Advancing Risk Assessment
http://www.nap.edu/catalog/12209.html

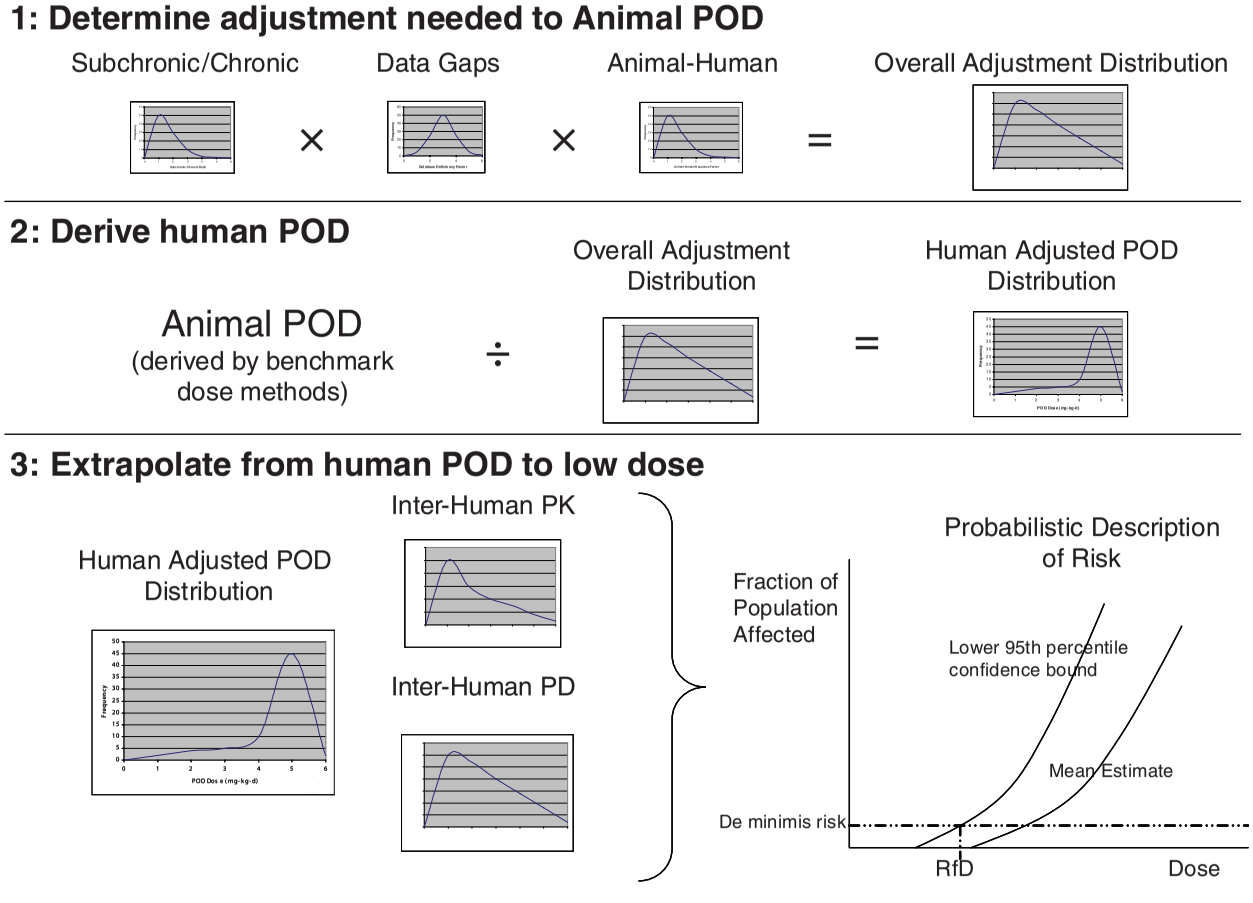

FIGURE 5-16 Steps in derivation of risk estimates for low-dose nonlinear end points. In Step 1 a distribution to adjust the animal POD to a human POD is constructed. Step 2 adjusts animal POD by this cross-species distribution. Step 3 uses human variability distribution to extrapolate from POD to lower risk. It results in probabilistic statement reflecting the proportion of human population adversely affected by exposure and uncertainty bounds on that estimate. It also enables identification of the RfD, that is, lower-bound estimate of dose that results in de minimis risk (for example, 10^{-5} of population is affected).

Copyright © National Academy of Sciences. All rights reserved.

Science and Decisions: Advancing Risk Assessment
http://www.nap.edu/catalog/12209.html

percentile of this distribution (that is, 98th percentile ≈ 2.01 × 2.17 × 2.17 = 9.5) (Baird et al. 1996; Hattis et al. 2002[3]).

- *Database-deficiency factor*. A dataset for 35 pesticides with "complete" toxicity-testing profiles was analyzed to compare reproductive, developmental, and chronic NOAELs (Evans and Baird 1998). It was possible to develop distributions for missing reproductive, developmental, or chronic toxicity data in terms of how much the POD can change by the addition of the missing data (Hattis et al. 2002). The data source is limited in terms of the type of chemicals assessed (pesticides) and the end points analyzed, but it provides an example of a useful approach to developing a distribution for this factor.
- *Animal-to-human extrapolation*. Cross-species differences in acute and subacute toxicity of anticancer drugs have been generalized to draw conclusions about animal-human differences in noncancer and cancer toxicity (Freireich et al. 1966; Travis and White 1988; Watanabe et al. 1992; Hattis et al. 2002). Animal-human interspecies distributions have been inferred from rat-mouse comparisons of cancer potency (Crouch and Wilson 1979), although because of the nature of the underlying data the distributions are likely to underpredict actual species differences. The results for cancer chemotherapeutic agents may have limited applicability. First, the agents are mostly direct-acting, so species differences in PK may not be as great for environmental chemicals. Second, the results are for a narrow range of end points (lethality and tolerated dose), and may not be representative of species differences for the more variable critical end points for environmental toxicants. Third, results are for acute and subacute exposures, and may not adequately represent cross-species differences for chronic exposures and more subtle end points. Indeed, Rhomberg and Wolff (1998) have shown that cross-species scaling observed with single-dose-lethal toxicity differs from the subacute toxicity. These authors hypothesize that "dose-scaling patterns across differently sized species should be different for single-dose and repeated-dose regimes of exposure, at least for severe toxic effects." The number of animal species studied is also an important consideration in developing the cross-species extrapolation distribution (Hattis et al. 2003). Further exploration of the issues raised is needed in developing interspecies distributions for application in EPA assessments.
- *Example derivation of the human POD*. In the examples given above, lognormal distributions replace uncertainty factors, and each factor is independent of the other. The overall adjustment is simple to calculate and does not have to be done numerically, using for example, Monte Carlo treatment. To obtain the human POD, the animal POD is divided by the overall adjustment factor, which for the sake of discussion is called here "$F_{A\rightarrow H\ POD\ adjust}$,"

$$\text{Human POD} = \text{Animal POD} \div F_{A\rightarrow H\ POD\ adjust}.$$

The overall adjustment is made up of three adjustments: for animal-to-human extrapolation, "$F_{A\rightarrow H}$"; for experiment duration from subchronic to chronic, "$F_{SC\rightarrow C}$"; and for data gaps, "F_{Gap}." Thus,

$$\text{Human POD} = (\text{Animal POD})/(F_{A\rightarrow H\ POD\ adjust}) = (\text{Animal POD})/(F_{A\rightarrow H} \times F_{SC\rightarrow C} \times F_{Gap}).$$

[3]A full version of this publication is available at http://www2.clarku.edu/faculty/dhattis. Updated results are published in Hattis and Lynch (2007).

Copyright © National Academy of Sciences. All rights reserved.

Science and Decisions: Advancing Risk Assessment
http://www.nap.edu/catalog/12209.html

If each factor is lognormally distributed, $F_{A \rightarrow H\ POD\ adjust}$ will be lognormally distributed; when a given adjustment is not needed its factor would be assigned a single value of 1.

The animal POD could be established as it is currently, or it could be described by the BMD distribution associated with its estimation. If the estimator of animal POD is lognormally distributed or is considered a constant, the human POD will be lognormally distributed. In this example, distribution of the animal POD is taken into account. Guidance on how a BMD should be defined for continuous outcomes would facilitate its current use as an animal POD (Gaylor et al. 1998; Sand et al. 2003) and for the probabilistic descriptions envisioned here.

The median value of the human POD distribution can be calculated by substituting the median values of the factors and the animal POD in the above equation.

$$\log (\text{Human POD}) = \log \text{Animal POD} - (\log F_{A \rightarrow H} + \log F_{SC \rightarrow C} + \log F_{Gap}).$$

For this case, each factor is assumed to be lognormally distributed,

$$\sigma^2_{\log \text{Human POD}} = \sigma^2_{\log \text{Animal POD}} + \sigma^2_{\log A \rightarrow H} + \sigma^2_{\log SC \rightarrow C} + \sigma^2_{\log Gap}.$$

The lower confidence bound on the human POD can be readily calculated. The human POD is the starting point for the extrapolation to lower doses based on information on human variability. A sample set of calculations is provided in Box 5-3 to illustrate how the above calculations can be made to derive a human POD.

Human Variability Distributions for Extrapolating from Human POD to Low Doses

- *Interindividual variability—PK Dimension*. Blood concentration information (AUC[4] and C_{max}[5]) were compiled for 471 data groups involving 37 drugs (Hattis et al. 2003) and summarized. A small number of data groups involved children under the age of 12 years. These PK data summaries that included young children (Ginsberg et al. 2002; Hattis et al. 2003) were then incorporated to yield a PK variability estimate for the overall population (Hattis and Lynch 2007). This work illustrates the feasibility of constructing PK variability distributions that are specific to particular age groups and clearance mechanisms. PK parameters have been derived from blood concentration data in children and adults and compiled according to type of agent, clearance pathway, or receptor (Ginsberg et al. 2002). Since these data come from a clinical setting in which the health of the studied subjects was impaired, and the characteristics of the treatment group may be similar, the data may not be representative of the general public. However, the researchers note the similarity of patterns of metabolizing-enzyme ontogeny in the databases and in vitro liver-bank specimens, suggesting that results from pharmaceutical studies may be generalizable.
- *Interindividual variability—PD Dimension*. From a database for 97 groups, Hattis et al. (2002) and Hattis and Lynch (2007) derived estimates of PD variability in (1) the chemical's reaching the target site after systemic absorption; (2) parameter change per delivered dose, the dose-response relationship at the active site (for example, beta-2-microglobulin spillage into urine in relation to urinary cadmium concentration); and (3) functional reserve,

[4]AUC is the area under the concentration-time curve that displays the complete time course of a chemical in a particular body compartment. AUC is sometimes used to represent the total dose in that compartment integrated over time.

[5]C_{max} is the maximum concentration of a chemical attained in a particular compartment after dosing.

Copyright © National Academy of Sciences. All rights reserved.

Science and Decisions: Advancing Risk Assessment
http://www.nap.edu/catalog/12209.html

BOX 5-3 Calculating a Risk-Specific Dose and Confidence Bound in Conceptual Model 2

I. Derivation of Human POD

$$\text{Human POD} = (\text{Animal POD})/F_{A\rightarrow H\ \text{POD adjust}} = (\text{Animal POD})/(F_{A\rightarrow H} \times F_{SC\rightarrow C} \times F_{Gap})$$
$$\log(\text{Human POD}) = (\log \text{Animal POD}) - (\log F_{A\rightarrow H} + \log F_{SC\rightarrow C} + \log F_{Gap})$$
$$\sigma^2_{\log F\ \text{Human POD}} = \sigma^2_{\log F\ \text{Animal POD}} + \sigma^2_{\log F\ A\rightarrow H} + \sigma^2_{\log F\ SC\rightarrow C}{}^2 + \sigma^2_{\log F\ \text{Gap}}$$

Assume:

Data gap is inconsequential:
$F_{Gap} = 1$, $\sigma_{\log F\ Gap} = 0$

Subchronic-to-chronic per Hattis et al. 2002:
50th percentile for $F_{SC\rightarrow C} = 2$, $\sigma_{\log F\ SC\rightarrow C} = \log[2.17] = 0.34$

Animal to human adjustment per Hattis et al. (2002) for sodium azide:
50th percentile for $F_{A\rightarrow H}$ 3.85, 95% upper bound 18.5, thus $\sigma_{\log A\rightarrow H} = \log(18.5/3.85)/1.645 = 0.42$ (Division by the 95% confidence bound is 1.645 standard deviations from the median in the standard normal distribution.)

Variability in animal POD:
lower 95% bound 2-fold difference from median; thus $\sigma_{\text{Animal POD}} = \log(2)/1.645 = 0.18$
$\Rightarrow$ Overall variability in human POD: $\sigma^2_{\text{Human POD}} = 0.34^2 + 0.18^2 + 0.42^2 = 0.32 = 0.57^2$

For animal POD (ED_{50}) of 1 mg/kg-d:
Human median POD (ED_{50}) = $1/(F_{A\rightarrow H}F_{SC\rightarrow C}F_{Gap}) = 1/(2 \times 3.85 \times 1) = 0.13$ mg/kg-d

Lower 95% confidence bound on human POD
$= (\text{median Human POD})/10^{[(1.645)(\sigma \log \text{Human POD})]} = 0.13/10^{[(1.645)(0.57)]} = 0.015$ mg/kg-d

II. Derivation of Risk-Specific Dose

Interindividual PK/PD variability (assume Hattis et al. 2002 distribution):
$\sigma_{logH} = 0.476$ (This estimate also is uncertain, with geometric standard deviation of 1.45)
The 10^{-5} individual is 4.25 standard deviations from the estimated human ED_{50}:
$10^{[(4.25)(0.476)]} = 105$

Median human dose with 10^{-5} risk:
(Median POD)/105 = 0.13/105 = 0.0012 mg/kg-d
Lower 95% bound on human dose with 10^{-5} risk: 0.006 µg/kg-d (This is calculated using a Monte Carlo procedure. It takes into account $\sigma_{\text{Human POD}}$ and the uncertainty in σ_{logH}.)

a factor inherent in many of the PD datasets but of which direct measurements in humans were not available. Hattis et al. (2002) took the first listed component as a component of PD rather than PK variability because it was related to reaching a specific organ, cell type, or subcellular constituent that is not typically addressed in physiologically-based pharmacokinetic models. The human interindividual variabilities derived for those components were combined to estimate the overall interhuman PD variability.

- *Overall distribution for human interindividual variability.* For the example here, overall human interindividual variability is described by a lognormal distribution with me-

Copyright © National Academy of Sciences. All rights reserved.

Science and Decisions: Advancing Risk Assessment
http://www.nap.edu/catalog/12209.html

dian of 1 and logarithmic (base 10) standard deviation σ_{logH}. Hattis et al. (2002) derived such a distribution for both PK and PD components from data for general systemic toxic effects on different agents, with a geometric standard deviation of 2.99 ($\sigma_{logH} = 0.476$; $10^{0.476} = 2.99$), indicating the median and upper 98th percentile human differ in sensitivity by a factor of 9. Human variability in response is chemical dependent. For some chemicals the difference between the median and 98th percentile is greater than a factor of 9, for others it will be less. Hattis and Lynch (2007) also describe the uncertainty in the variability estimate. The estimate of 0.476 for σ_{logH} has its own geometric standard deviation of 1.45. Because these characterizations of variability are limited by the relatively small numbers upon which the estimates are based, this uncertainty estimate may have a downward bias.

Calculation of Risk-Specific Dose and Confidence Bound

A distribution of human variability can be applied to move from the human POD down the dose-response curve, as illustrated in the set of calculations in Box 5-3. These calculations illustrate a generic case with an animal median ED_{50} value of 1 mg/kg-day.

In Box 5-3, as done by Hattis et al. (2002) and (Evans et al. 2001), the ED_{50} was chosen as the POD. Because the ED_{50} is at the center of the animal dose-response curve, there is less uncertainty in its measurement, and it is not as heavily influenced by interanimal variability as a response at the tail of the distribution might be. In addition, in many animal experimental datasets, the ED_{50} is not likely to be as influenced by the dose-response model selected to analyze the data relative to other effect levels. But there are other factors, such as intra-individual variability and the extent that this may play a role in the dose-response relationship. Any implementation of this approach by EPA would have to develop a process for selecting the POD for risk extrapolation for nonlinear end points.

Interhuman PK and PD distributions would ideally be derived with chemical-specific data on the differences possible among human populations. However, this type of information is usually lacking. Therefore, generic distributions based on surrogate chemicals and end points will be needed. Specific distributions for related chemicals and end points of interest may be possible. The first tier of a default distribution may be one built on a broad array of structurally dissimilar chemicals tested in different types of systems (from in vitro to in vivo) for different end points. The Hattis et al. (2002) effort to collect and analyze mostly clinical human data is a good initial effort at characterizing human PD variability. However, an important consideration with regard to this and related exercises is whether they fully capture PD variability, given the limited array of data studied. Data on small numbers of people may be a useful beginning but provide little information on overall interhuman variability. Even when multiple studies are combined so that data on greater numbers of people are tabulated, they still might not capture the broad spectrum of PD variability caused by differences in age, genetics, diet, health status, medications, and exposure to other agents.

Greater relevance may be achieved by applying PD variability information on prototypical chemicals in the same class as the chemical of interest. When there is a much larger and substantial database on one particular toxicant in a structural series, there is the potential to apply the information to others in the series on the basis of relative-potency approaches, as described in Chapter 6. A similar analogy may also be useful for assessing interhuman PD variability if the toxicity end points of the prototype and of the chemical of interest match well. For example, human variability in the renal response to cadmium, as assessed on the basis of beta-2-microglobulin leakage, may be relevant to other heavy metals, such as mercury and uranium, that can also damage the kidney (Kobayashi et al. 2006). Another possibility is that the degree of interhuman variability can be gleaned from studies of

Copyright © National Academy of Sciences. All rights reserved.

Science and Decisions: Advancing Risk Assessment
http://www.nap.edu/catalog/12209.html

environmental mixtures to which populations are exposed. Biomarkers of exposure—such as urinary 1-hydroxypyrene, a marker of exposure to polycyclic aromatic hydrocarbons (PAHs)—can be related to biomarkers of effective internal dose (such as bulky DNA adducts and urine mutagenicity) and effect (such as chromosomal damage in peripheral lymphocytes). Evaluations of such markers in coke-oven workers, bus drivers, and the general population ingesting charcoal-broiled meat or inhaling cigarette smoke provide a database from which interindividual variability in response to carcinogenic PAHs may be deduced (Santella et al. 1993; Kang et al. 1995; Autrup et al. 1999; Siwinska et al. 2004).

Thus, the data gap represented by interhuman PD variability presents a critical research need that can be approached by mining the existing epidemiology literature and by designing new studies in which biomarkers of exposure and effect are used to describe variability in sensitivity to health outcomes in similarly exposed people.

There are likely to be a number of cases in which the approach illustrated above can be used to derive an RfD. Sometimes, however, there will be a well-defined sensitive subgroup. The RfD for the pesticide alachlor is based on hemolytic anemia in dogs (EPA 1993); the background incidence of hemolytic anemia in humans is generally very low except in ethnic groups in which, because of inherited traits (such as glucose-6-phosphate dehydrogenase deficiency), the risk is higher (Sackey 1999). In cases like this one, an analysis focusing on describing risks to the sensitive subgroups would be needed (see Figure 5-12).

Conceptual Model 3: Low-Dose Linear Individual and Population Dose-Response Relationship

Here linear dose-response processes govern the dose-response relationship for individuals, as may occur for cancer and other complex toxic processes, and consequently the population dose response relationship is low-dose linear. This is unlike the previous two conceptual models, which described population dose-response distributions that arise when the dose-response relationship in an individual has a threshold. A possible approach to default analysis following this conceptual model is presented below. It emphasizes probabilistic descriptions of the uncertainty in the dose-response relationship and descriptions of variability among individuals exposed to the same dose.

Approach

This approach to dose-response analysis begins, as do the other examples above, with the derivation of the human POD distribution. When derived from animal data, the human POD is based on the animal POD and distributions of adjustment factor, such as for interspecies differences and study duration less than a lifetime. Here, the POD is taken from a model fitted at a dose in the lower end of the observable response range, and does not use an ED_{50}. Risk at lower dose than this POD for the median person is estimated by linear extrapolation, that is, risk is assumed to decrease linearly with dose below the POD. However, as illustrated in Chapter 4 (Table 4-1), people exposed to the same dose will differ in risk. Estimates of the spectrum of individual risks at a given dose can be based on a distribution that describes interhuman variability. The individual dose-response relationships allow the calculation of the population dose-response curve. This approach to dose-response assessment is illustrated in Figure 5-17 and through the case study for 4-aminobiphenyl.

Copyright © National Academy of Sciences. All rights reserved.

Science and Decisions: Advancing Risk Assessment
http://www.nap.edu/catalog/12209.html

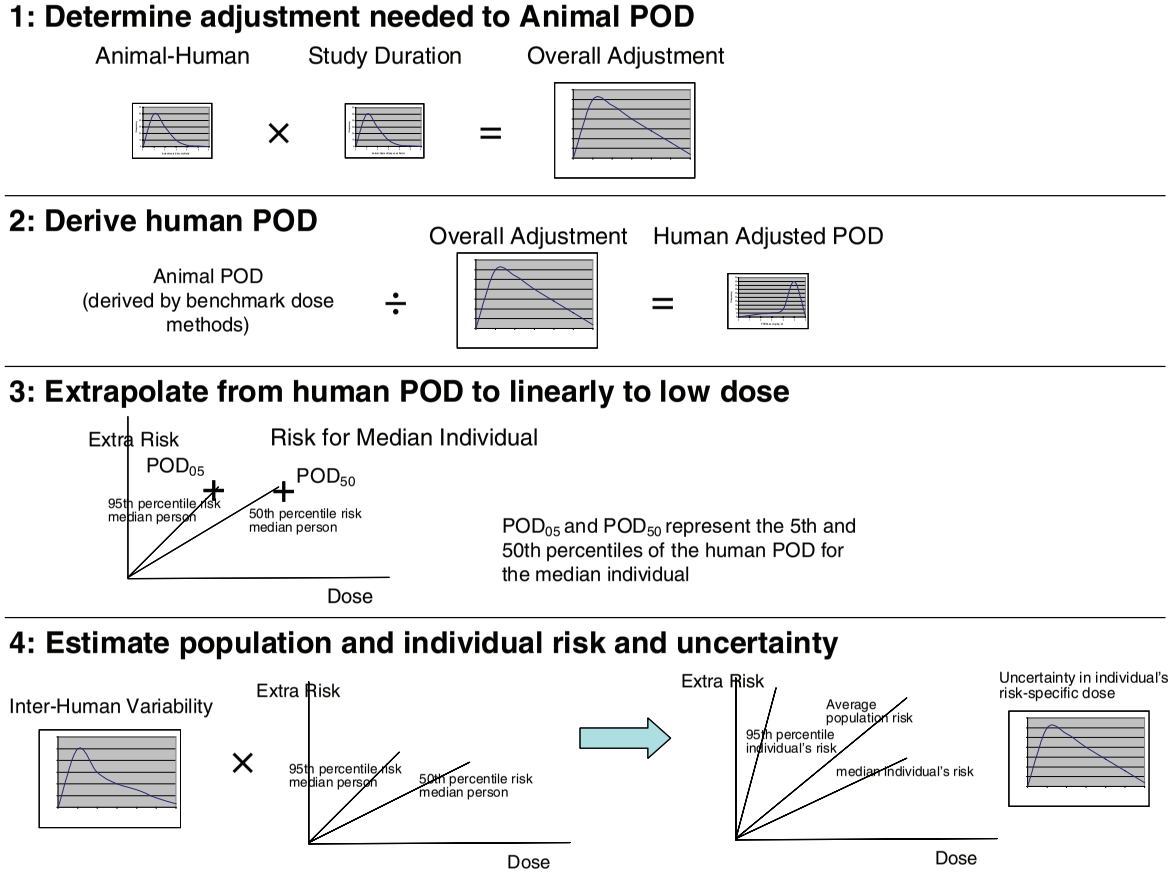

FIGURE 5-17 Steps to derive population and individual risk estimates, with uncertainty in estimates from animal data. Step 1 involves derivation of adjustment distribution to convert animal POD to human POD. Step 2 involves derivation of human POD from this distribution. Step 3 is linear extrapolation from POD to lower doses for median person. Lower bound on human POD is used to derive upper-bound risks for median person. Step 4 involves applying interindividual variability distribution to estimate average risk to population and risks to individuals with different degrees of sensitivity, with uncertainty in estimates.

Copyright © National Academy of Sciences. All rights reserved.

Science and Decisions: Advancing Risk Assessment
http://www.nap.edu/catalog/12209.html

Implications of the Approach

Functionally, the approach would change dose-response characterizations for low-dose linear carcinogens in two basic ways. First, there would be an explicit characterization of uncertainty in the human POD that accounted for uncertainty in the cross-species extrapolation and the statistical fit to the dose-response data. EPA could choose to report particular percentile values, such as the upper 95th percentile. EPA could describe the population excess cancer risk associated with dose D as the plausible upper bound of the excess risk, taking into account uncertainty in the population dose-response relationship and variability in the individual dose-response relationship. The excess-risk estimate for a person whose susceptibility puts him or her at the 95th or some other percentile of the population could also be separately reported.

Second, when the underlying variability distributions are right-skewed, as in the case of the lognormal distribution, the population risk estimate emerging from the analysis will be greater than the estimate for the median individual. The mean or "expected value" will exceed the median value by some amount that depends on the assumed shape of the distribution of interindividual variability in susceptibility.

Recommended Default for Interindividual Variability in Cancer Susceptibility

An assumption that the distribution is lognormal is reasonable, as is an assumption of a difference of a factor of 10-50 between median and upper 95th percentile people, as indicated by the series of examples provided in Chapter 4. It is clear that the difference is significantly greater than a factor of 1, the current implicit assumption in cancer risk assessment. In the absence of further research leading to more accurate distributional values or chemical-specific information, the committee recommends that EPA adopt a default distribution or fixed adjustment value for use in cancer risk assessment. A factor of 25 would be a reasonable default value to assume as a ratio between the median and upper 95th percentile persons' cancer sensitivity for the low-dose linear case, as would be a default lognormal distribution. A factor of twenty-five could be interpreted as a factor of 10 for pharmacokinetic variability, and a factor of 2.5 for pharmacodynamic variability. For some chemicals, as in the 4-aminobiphenyl case study below, variability due to interindividual PK differences can be greater. In a cancer process, with long latency and multiple determinants, PD variability could be considerably greater than the suggested default. PD differences would include the various degrees among people in DNA repair and misrepair, surveillance of mutated cells, and accumulation of additional mutations and other factors involved in progression to malignancy.

A common assumption for noncancer end points is an overall factor of 10 to account for interindividual variability—3.2 or 4 uncertainty factor for PK differences and 3.2 or 2.5 for PD differences (EPA 2002a; IPCS 2005). For genotoxic metabolically activated carcinogens, Hattis and Barlow (1996), considering activation, detoxification and DNA repair alone, found greater PK variability with individuals at the median and the 95th percentile differing by a factor of 10. The factor was a central estimate, some chemicals exhibited greater and others lesser PK variability. In the 4-aminobiphenyl case discussed below, additional physiologic factors such as storage in the bladder contributed to human variability in PK elements.

The suggested default of 25 will have the effect of increasing the population risk (average risk) relative to the median person's risk by a factor of 6.8: For a lognormal distribution, the mean to median ratio is equal to $\exp(\sigma^2/2)$. When the 95th percentile to median ratio is 25,

Copyright © National Academy of Sciences. All rights reserved.

Science and Decisions: Advancing Risk Assessment
http://www.nap.edu/catalog/12209.html

σ is 1.96 [=ln(25)/1.645], and the mean exceeds the median by a factor of 6.8. If the risk to the median human were estimated to be 10^{-6}, and a population of one-million persons were exposed, the expected number of cases of cancer would be 6.8 rather than 1.0.

Thus under this new default, the value for the median person would remain as provided by the current approach to cancer risk assessment; for a default of a factor of 25, the average would be higher by a factor of 6.8. It would be important for the cancer risk assessment to express interindividual variability by showing the median and average population risks, as well as the range of individual risks for risk-management consideration.

Case Study: 4-Aminobiphenyl

4-Aminobiphenyl is a known cause of human bladder cancer. It was once used as a dye intermediate and rubber antioxidant, but its use was curtailed after findings of bladder cancer in substantial numbers of workers. Current exposures are due mostly to cigarette-smoking, which increases bladder-cancer risk by 2-10 times. The compound binds to bladder DNA and is mutagenic in a variety of test systems, including human cell culture. It has the hallmarks of low-dose linearity and is implicated as a cause of bladder cancer in smokers exposed to relatively low doses and quite recently in female never-smokers in Los Angeles County exposed to environmental tobacco smoke (Jiang et al. 2007). The compound has been extensively studied and found to have marked interindividual differences in activation and detoxification, and higher risk has been observed in slow acetylators, who detoxify it less efficiently (Gu et al. 2005, Inatomi et al. 1999). It is presented to illustrate how human interindividual variability can be addressed in dose-response assessment when reasonably high-quality data are available.

Estimating Variability in Human Susceptibility to 4-Aminobiphenyl

Bois et al. (1995) modeled interindividual heterogeneity in human cancer risk using data on differences among humans in their PK and physiologic handling of 4-aminobiphenyl. Briefly, the compound is thought to be activated via *N*-hydroxylation by CYP1A2, although recently other enzymes have also been found to be involved (Tsuneoka et al. 2003; Nakajima et al. 2006). A major detoxification pathway is *N*-acetylation. To simulate interindividual variability in pharmacokinetics, parameters describing the absorption, distribution, activation, detoxification, and urinary excretion were varied according to human ranges found in the literature. Distributions of the formation of the proximate carcinogen and its binding to urinary-bladder DNA were simulated. The latter can be used to describe possible differences in susceptibility due to physiologic and PK factors and is shown in Figure 5-18.

The DNA-binding distribution accounts for human differences only up to the point of binding and does not address PD differences. The DNA-binding distribution therefore can be considered an undercharacterization of overall human variability. The upper and lower bounds for the PK-based distributions shown in Figure 5-18 differed from the geometric mean by factors of 16 and 26, respectively. The distribution of human interindividual variability would be greater than indicated by the PK-based distributions because of PD differences among people.

For the 4-aminobiphenyl case study an estimate of interindividual variability of a range of 50 (ratio of 95th percentile to median person) is assumed for the purposes of illustrating the incorporation of variability into cancer dose-response modeling. It reflects the factor of roughly 20-30 between median and upper 95th percentile individual sensitivity in pharmacokinetics and a modest factor for variability factors pertinent to PD differences in carci-

Copyright © National Academy of Sciences. All rights reserved.

Science and Decisions: Advancing Risk Assessment
http://www.nap.edu/catalog/12209.html

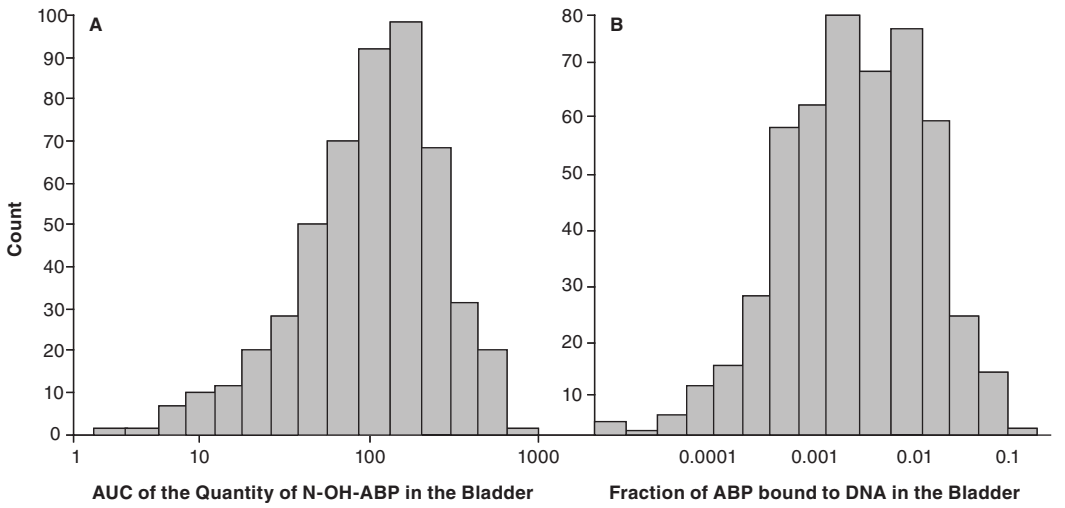

FIGURE 5-18 *A*, AUC for proximate carcinogen in bladder in units of nanograms-minutes simulated for 500 people. *B*, simulated fraction bound in bladder, presumed to indicate differences in susceptibility due to PK and physiologic parameters. Upper 95% confidence limit is factor of 16 above geometric mean of 0.0034 and factor of 26 above lower confidence limit. Source: Bois et al. 1995. Reprinted with permission; copyright 1995, *Risk Analysis*.

nogenesis. As noted above, this may be an underestimate, the range of 50, assumed in the calculations for this case study, corresponds to a geometric standard deviation of 10.8 and a standard deviation in natural log space of 2.38 (σ_{lnH}), and in base 10 log space of 1.03.

Derivation of Median Human POD and Slope for 4-Aminobiphenyl

Despite the known causal association between human bladder-cancer risk and 4-aminobiphenyl, human exposure estimates in occupational studies may be insufficient for establishing reliable dose-response relationships, and the assessment may have to be based on animal data, as is done here. Fitting dose-response models to a sensitive site observed in the animals—liver tumors in female mice exposed by gavage—results in an ED_{10} of 0.1 mg/kg-d with a lower 95% confidence bound of 0.070 mg/kg-d. That corresponds roughly to a $\sigma_{\text{logAnimal POD}}$ of 0.09, assuming a lognormal uncertainty distribution. For cross-species extrapolation to adjust the animal POD to the human POD, doses are assumed to have equal effectiveness if the human dose is reduced consistently with three-fourths bodyweight scaling. As described above, data are available on acute and subacute toxicity in different species. Hattis et al. (2002) derived an uncertainty estimate of 0.416 for $\sigma_{\text{logA}\rightarrow\text{H}}$ from those data, and also found that an additional small factor slightly increasing the uncertainty estimate was merited. However, for cancer end points, which result from more protracted and complex biologic processes, that value can be presumed to substantially understate the actual uncertainty. Nonetheless, it is adopted for the illustrative example here with the recognition that the estimate may be low. The median estimate for cross-adjustment scaling would be 7.3 [that is, $(70/0.025)^{(1-0.75)}$, assuming bodyweights of 70 kg and 0.025 kg for humans and mice, respectively]. Thus, the median human POD would be 0.014 mg/kg-d (0.1/7.3), and the slope of the dose-response curve at the POD would be 7.5 $(\text{mg/kg-d})^{-1}$ [–ln(0.9)/0.014].

Copyright © National Academy of Sciences. All rights reserved.

Science and Decisions: Advancing Risk Assessment
http://www.nap.edu/catalog/12209.html

The confidence interval would take into account uncertainty resulting from fitting the dose-response model to the data, the cross-species extrapolation, and other factors. For the sake of illustration, the first two factors are accounted for here, and the resulting $\sigma_{humanPOD}$ is 0.43 $[(0.09^2 + 0.42^2)^{1/2}]$, reflecting a lower 95% bound on the median human POD of 0.003 mg/kg-d ($0.014/10^{1.645\sigma} = 0.014/5.1$) and an upper 95% bound of 38 $(mg/kg\text{-}d)^{-1}$ [(7.5)(5.1)].

Derivation of Individual and Population Dose-Response Relationships for 4-Aminobiphenyl

On the basis of the interindividual variability estimate noted above, the population and individual percentile dose-response relationships can be estimated with the uncertainty estimates for those functions. The slope of the population dose-response curve can be calculated from the risk, averaged among individuals, at a given dose. For a lognormally distributed variable, the derivation of the mean, μ, from the median involves a simple calculation [μ = (median)(exp{$\sigma^2/2$}), where σ is expressed in base e units and "exp{$\sigma^2/2$}" represents base e to the power {$\sigma^2/2$}]; for this case, with the human interindividual variability estimate, σ_{lnH} = 2.38, the mean potency is 126 $(mg/kg\text{-}d)^{-1}$ [(7.5)(exp{$2.38^2/2$}) = (7.5)(16.9)]. At low doses, population risk is calculated by multiplying the population potency by the dose. The uncertainty bound on this estimate is derived by considering the uncertainty in the adjustment factors and in the model fit at the POD.

At low doses, the risk for the 95th percentile person is given by multiplying the dose by 376 $(mg/kg\text{-}d)^{-1}$ [376 = (7.5)(50)], and the dose associated with a 10^{-5} risk for the 95th percentile sensitive person would be 3 µg/kg (that is, [14 µg/kg]/50). The uncertainty bounds around this dose estimate would be given by the human POD distribution, represented by $\sigma_{humanPOD}$. The lower confidence bound on the estimate for this person would be determined by σ as described above (for example, the 95% lower bound would be [(3 µg/kg-d)/$10^{1.645\sigma_{human\ POD}}$ = 0.6 µg/kg-d]). This example does not capture all sources of uncertainty and is provided only to illustrate an approach.

Mathematical Framework for Conceptual Model 3

Human low-dose risk[6] from a given dose D of toxicant could be expressed as

$$Risk_H = Slope_H D = (Slope_{BMD} F_{H-A}) D. \quad (1)$$

$Risk_H$ here is the incremental increase in risk above background, also called "extra risk." In current practice, $Slope_{BMD}$ is the slope[7] of the dose-response curve at the BMD. The cross-species factor, $F_{H\text{-}A}$, adjusts for differences in effect in humans compared with animals exposed to the same dose and is usually greater than 1. As discussed above, $F_{H\text{-}A}$ is typically expressed as two factors: one to account for human-animal differences in pharmacokinet-

[6]If a quantal linear-regression model is fitted and the risk over the dose-response range (π_D) can be given by $\pi_D = 1 - \exp(-\beta_0 - \beta_1 D)$. Extra risk (ER) can be defined as: $ER(D) = (\pi_D - \pi_0)/(1 - \pi_0)$. This model reduces to $ER(D) = 1 - \exp(-\beta_1 D)$. For a specified benchmark response (BMR), the BMD is defined in this model as $BMD = -\ln(1 - BMR)/\beta_1$. When the relationship between extra risk and dose is quadratic, $\pi_D = 1 - \exp(-\beta_0 - \beta_1 D - \beta_2 D^2)$.

[7]The $Slope_{BMD}$ could be defined as the slope of the line tangent to the ER(D) curve at D = BMD, that is, $Slope_{BMD}$ = ER(BMD) = d/dD[ER(D)] evaluated at D = BMD. For ER(D) defined in the context of a quantal linear model, this reduces to $Slope_{BMD} = ER'(BMD) = \beta_1 \exp(-\beta_1 BMD)$ evaluated at the estimated BMD. For simplicity and transparency, however, the following approximation can be used: $Slope_{BMD}$ = BMR/BMD, which corresponds to the slope of the line connecting (BMD, BMR) and (0,0).

Copyright © National Academy of Sciences. All rights reserved.

Science and Decisions: Advancing Risk Assessment
http://www.nap.edu/catalog/12209.html

ics and the other to account for human-animal differences in pharmacodynamics ($F_{H\text{-}A} = F_{H\text{-}A\ PK}F_{H\text{-}A\ PD}$). It is a means of converting the animal slope (for example, in $[mg/kg\text{-}d]^{-1}$) to a human slope and can be thought of as going from the median animal to the median human. In cases in which cross-species differences in pharmacokinetics were used to derive the $Slope_{BMD}$, $F_{H\text{-}A}$ would be represented by $F_{H\text{-}A\ PD}$.

Each factor in Equation 1 may represent a model, a single number, or a distribution, depending on the nature of the data and the goal of the analysis. Variability in exposure could also be incorporated through some distribution of exposures (for example, as $D \sim G_D()$). Uncertainty in dose extrapolation or in the animal-human extrapolation could be addressed by $F_{H\text{-}A}$ as distributions (for example, $F_{H\text{-}A} \sim G_{H\text{-}A}()$). It is important to distinguish variability in risk among individuals—that is, the difference in risk from individual to individual—from uncertainty, which describes our lack of knowledge of the risk. The goal here is to enable such expressions as "The risk of effect does not exceed *x* for the *y*th percentile individual, stated with a confidence interval of *z1- z2*." The 4-aminobiphenyl case provided an example of how that might be done.

In some cases, as a default, it may be convenient and appropriate to describe uncertainty in $Risk_H$ mathematically with a lognormal distribution, for example, if the uncertainty in each factor in Equation 1 can be represented by a lognormal distribution. In this case, Equation 1 may be re-expressed as

$$\text{Log Risk} = \log Slope_{BMD} + \log F_{H\text{-}A} + \log D.$$

For this simplistic case,

$$\sigma^2 = \sigma_{logSLOPE}^2 + \sigma_{logF}^2 + \sigma_{logD}^2. \quad (2)$$

To the extent that σ_{logD} represents differences in exposure rather than uncertainty, it would not be incorporated as above but tracked separately to be combined with the human susceptibility distributions described below. The 4-aminobiphenyl case illustrates how variability in PK factors may lead to considerably greater risks in some people than others and how this might be taken into account quantitatively. Formally introducing human PD variability into mathematical descriptions is more challenging, and in the case example below a default distribution is assumed. The risk for the *y*th percentile person may be described by

$$Risk_{H\ yth} = Slope_{BMD}F_{H\text{-}A}DV_{H\ yth}, \quad (3)$$

where V_{Hyth} is the *y*th quantile of the distribution that describes the ratio of the *y*th percentile person to the median person. If the uncertainty in V_{Hyth} and the other elements of the uncertainty are described by a lognormal distribution, the overall uncertainty represented by σ^2 would be described by adding a term.

σ_{logV}^2, to the terms given in Equation 2.

Multiple Dose Dependent Modes of Action

The most recent EPA (2001c) dose-response assessment for chloroform, the drinking-water disinfection byproduct, assumed that sustained or repeated cytotoxicity followed by regenerative hyperplasia was probably the cause of kidney and liver cancer observed in rodent bioassays. A margin-of-exposure (MOE) approach was recommended for the evalua-

Copyright © National Academy of Sciences. All rights reserved.

Science and Decisions: Advancing Risk Assessment
http://www.nap.edu/catalog/12209.html

tion of carcinogenic exposures to the compound. However, the EPASAB (2000, p. 12) noted there is "some possibility that genotoxicity could contribute to the dose-response at low doses" for the observed kidney tumors and called for the agency to address the general issue of mixed modes of action by "beginning to develop a reasonable means of estimating the most likely and upper bound estimate of potential contribution of a 'genotoxic' component to the carcinogenic activity."

Dose-response analysis of chemicals whose end points are associated with multiple MOAs is challenging. The EPA (2005b) *Guidelines for Carcinogen Risk Assessment* state (p. 3-22) that

> if there are multiple modes of action at a single tumor site, one linear and another nonlinear, then both approaches are used to decouple and consider the respective contributions of each mode of action in different dose ranges. For example, an agent can act predominantly through cytotoxicity at high doses and through mutagenicity at lower doses where cytotoxicity does not occur. Modeling to a low response level can be useful for estimating the response at doses where the high-dose mode of action would be less important.

Although that may have been the case for chloroform, the agency decided to take a low-dose nonlinear approach to characterize the risks associated with the chemical, and applied noncancer RfD methodology. In cases like that of chloroform, the slope at high doses would not give a good indication of the low-dose slope. For cases with low-dose linearity in an individual's response in which the high-dose response may be significantly influenced by a nonlinear MOA neither conceptual model 2 nor projection of low-dose risk from a high-dose BMD is satisfactory. In such cases an alternative default approach is suggested.

At low doses, the linear MOA can be expected to dominate. A modifying factor, M_S, could account for the change in slope. The adjustment factor would be based on mechanistic understanding. In this case risk (that is, "extra risk" as defined above) would be given by

$$\text{Risk}_H = [\text{Slope}_H]D = [\text{Slope}_{BMD} M_S F_{H\text{-}A}]D, \quad (4)$$

where the terms Slope_{BMD}, $F_{H\text{-}A}$, and D are as defined above for Equation 1, with Slope_{BMD} estimated as described above.

For cases like 4-aminobiphenyl, M_S would have a value of 1. For cases like chloroform, it would have a value less than 1 and would probably be the subject of controversy and debate. Nonetheless. this M_s provides a vehicle for addressing potential low-dose linearity in cases in which there is strong evidence that the slope observed at high doses overpredicts the low-dose slope.

M would serve the same purpose as the "dose and dose-rate effectiveness factor" adopted to adjust the slope of the dose-response curve observed at relatively high doses to predict radiation risk at low doses (ICRP 1991; EPA 1998; NRC 1998b; ICRP 2006; NRC 2006c; Wrixon 2008). Multiple mechanisms of toxicity may exist for a single agent, some of these mechanisms may have nonlinear-dose-response characteristics, or so-called dose-dependent transitions (Slikker et al. 2004). In considering values for M, any dose-dependent transitions would be considered in the context of background exposures and disease processes affecting these toxicity mechanisms. The selection of M would be a science policy call.

IMPLEMENTATION

The committee recognizes that the unified framework introduces additional needs for data and analyses into the risk-assessment process. The data and analyses may take time to

Copyright © National Academy of Sciences. All rights reserved.

Science and Decisions: Advancing Risk Assessment
http://www.nap.edu/catalog/12209.html

develop, and development of an implementation strategy will be important. The committee notes that the framework can be implemented in the short term by establishing default distributions. For noncancer end points, the defaults will enable a probabilistic basis of establishing the RfD and characterizing noncancer risks; for cancer-risk characterization, they will enable incorporation of interhuman variability. Use of default distributions for adjustments in extrapolations, rather than default point-estimate uncertainty factors, provides an improved representation of variability and uncertainty and offers an opportunity for further refinements and incentives to gather and analyze existing information and to generate new data targeted to specific extrapolation needs. As experience accrues, guidelines will also be important to aid in the application of the defaults and to ensure consistency in the implementation of the framework. In the development of guidelines, the committee encourages attentiveness to issues regarding the use of defaults addressed in Chapter 6 and has concerns about the approach taken to ascertain a mutagenic MOA for genotoxic carcinogens (see discussion in the Mode-of-Action Assessment section above) in application of the guidelines to address early-life sensitivity to cancer (EPA 2005c). The committee has illustrated the ideas advocated in this chapter with conceptual models and example calculations. Assumptions and simplifications are used to make the examples tractable and clear, not to prescribe any particular approach or value.

Table 5-1 summarizes major aspects of the unified framework in terms of data needs, potential utility of defaults as interim placeholders for better-researched and better-defined distributions, and implementation. A number of other sources of uncertainty and variability that often arise in dose-response assessment are not peculiar to the proposed unified framework and so are not addressed in the table; some of these issues and their associated default approaches are described in Chapter 6.

An implementation plan can be devised to phase in the unified framework. Some considerations and suggestions for developing the plan are presented in Table 5-1. Default distributions can initially be based on datasets that can be augmented with adjustments or other distributional assumptions to account for inferences that generalize from small numbers of people, of chemical case studies, and of end points to large populations, numbers of chemicals, and numbers of effects. As more data are collected and variability is better understood, the uncertainty portion of the default distribution may decrease. Emerging technologies, such as toxicogenomics and high-throughput assays, will highlight pathways that are at the crossroads of disease causation and toxicant action and will assist in the incorporation of background additivity and variability components. The implementation plan should be associated with a research agenda that will, over time, enable refinement of distributional approaches to dose-response assessment. Finally, EPA guidance will clearly be needed in order to implement the unified framework, including conduct of the background exposure and vulnerability assessments, departure from the linear default, establishment of distributions for the analysis, model selection, and so on. The development and roll out of guidelines and policies will be an essential component of any implementation plan, as well as ample opportunity for stakeholder involvement, scientific peer review and mid-course correction to address false starts and mis-steps.

Copyright © National Academy of Sciences. All rights reserved.

Science and Decisions: Advancing Risk Assessment
http://www.nap.edu/catalog/12209.html

TABLE 5-1 Potential Approaches to Establish Defaults to Implement the Unified Framework for Dose-Response Assessment

Analytic Step	Data Need	Testing and Implementation Issues	Potential Approach for Establishing Defaults in the Near Term
Cross-species extrapolation	Relative sensitivity to toxicants, comparing rodent with human	Moving from default point estimate to distribution adds complexity and encounters data limitations; literature on acute and subacute effects and direct-acting drugs is used mostly in comparisons, and small numbers of people studied may not be representative of human population	Base default distribution on wide sample of drugs and toxicants for which there are data on rodents and humans (drug trials, clinical toxicologic and epidemiologic studies) for similar end points and on adjustments to address data gaps; look to specify distributions to particular classes of chemicals and comparisons of particular rodents (mouse vs rat vs hamster) with humans; consider using bodyweight scaling for PK portion of extrapolation if overall distribution covering PK and PD cannot be derived; develop default distribution to describe uncertainty in bodyweight scaling
Interindividual PK variability in humans	PK differences among life stages, disease states, genetic polymorphisms, drug interactions	PK datasets on susceptible groups (such as children) are difficult to obtain; default may have to be based primarily on drug literature, which is also limited	Derive default distribution of PK variability based on analogy with drug literature and, to extent possible, made spceific to particular enzyme pathways, types of receptors, and classes of chemicals; use PBPK Bayesian and Monte Carlo approaches to evaluate implications of variability in enzyme pathways for overall PK variability; consider adjustments to address small samples and other biases in derivation
Interindividual PD variability in humans	PD differences in population with respect to various types of end points, including cancer	Human PD response is likely to vary widely, especially in groups that are difficult to study (such as children, elderly); it is unclear how to consider and integrate clinical, precursor, and other upstream end points and how to separate PK from PD variability	Base default distribution on broad array of human responses, chemicals, and end points from drug testing and high-quality epidemiologic studies; use information on background exposures and vulnerabilities to develop default; develop distributions specific to chemical classes, end points (such as cancer, endocrine, and acute toxicity), and humans to extent possible; consider adjustments to address small samples and other biases in derivation

Copyright © National Academy of Sciences. All rights reserved.

Science and Decisions: Advancing Risk Assessment
http://www.nap.edu/catalog/12209.html

TABLE 5-1 Continued

Analytic Step	Data Need	Testing and Implementation Issues	Potential Approach for Establishing Defaults in the Near Term
Background exposures	Low-dose interaction studies for chemicals with similar MOA	Human population has numerous background exposures; MOAs are difficult to define; when they are defined, interaction with other chemicals can be difficult to predict at different doses and dose ratios and in different species, ages, and organs; mechanistic and interaction data are limited	Develop guidance to judge whether background exposures (and vulnerability) are sufficiently unimportant to reject linearity at low doses; when it is rejected, use probabilistic approach to develop RfD, using interindividual variability and other distributions
Background vulnerability[1]	Sensitive epidemiologic and mechanistic studies relating chemical exposures and disease processes; biomonitoring data	Human population has numerous degenerative and disease processes; it is difficult to sort relevance to particular MOA; data on chemical-disease interaction are insufficient	Establish guidance to judge whether background vulnerability, conditions, and exposures are sufficiently unimportant to reject linearity at low doses; when it is rejected, use probabilistic approach to develop RfD, using interindividual variability and other distributions
Low-dose extrapolation defaults	MOA information defining chemical effect at target and interaction with background processes	It is difficult to obtain low-dose data in relevant test systems; chemicals can have mixed MOAs; different models can fit high-dose data equally but differ at low dose	Continue assumption that carcinogens have low-dose linear response unless sufficient data support other approaches; develop guidance for noncancer low-dose response and linear extrapolation due to background additivity and vulnerability (conceptual model 1); formally adopt assumption that genotoxic chemicals (clastogens, mutagens) cause cancer via a mutagenic MOA
Low-dose linear slope factor—M adjustment	Dose-response data over wide dose ranges in human and animal studies and related mechanistic data	Data from epidemiologic and toxicologic studies are limited; there is need to know how to use biologic models in considering mechanistic data	Develop series of default M factors based on mechanistic considerations and human and animal observations to apply in different situations (such as saturation phenomena or high-dose cytotoxicity that influences carcinogenicity of chemicals with some genotoxic activity)

[a]Susceptibility to endogenous (for example, age, gender, genetics, pre-existing health deficits and disease) and exogenous factors (exposure to agents) and due to variability in exposure.

Copyright © National Academy of Sciences. All rights reserved.

Science and Decisions: Advancing Risk Assessment
http://www.nap.edu/catalog/12209.html

CONCLUSIONS AND RECOMMENDATIONS

Conclusions

This chapter reviews the current paradigm for characterizing the dose-response relationships of compounds for both cancer and noncancer end points and supports the following conclusions:

- Separation of cancer and noncancer outcomes in dose-response analysis is artificial because noncancer end points can occur without a threshold or low-dose nonlinearity on the population level and in some cases on the individual level. Similarly, the MOA for carcinogens varies and requires a flexible but consistent analytic framework. The separation not only is scientifically unjustified but leads to undesirable risk-management outcomes, including inadequate attention to noncancer end points, especially in benefit-cost analyses.
- The current formulation of the RfD is problematic because of its application as a determinant of risk vs no risk of regulatory importance, and it lacks a quantitative description of the risk at different doses. It hinders risk-risk and risk-benefit comparisons and risk-management decision-making and does not make the best possible use of available scientific evidence.
- Cancer risk assessment typically lacks a quantitative description of interindividual variability. That leads to an incomplete description of the range of risk possible in the population. Noncancer risk assessment addresses interindividual variability, but cancer risk assessment typically does not; this reflects the implicit default assumption that human cancer susceptibility does not vary (see Chapters 4 and 6). The argument that the linear dose-response extrapolation procedure covers the omission (EPA 2005b) is unsupported and presents a separate consideration that should not be confused with the need to describe risk differences among individuals in addition to high-dose–low-dose extrapolation. The approach adopted in the current carcinogen guidelines (EPA 2005b) that considers variability only when a sensitive subpopulation can be identified for a particular chemical is limited by a lack of chemical-specific data. It also ignores the appreciable scientific knowledge of human interindividual variability in sensitivity (see, for example, Table 4-1), which can form the basis of general assumptions regarding variability when chemical-specific data are absent. The supplemental guidance regarding children (EPA 2005c) is an important step in the right direction, but variability in the general population should also be addressed.
- Uncertainty factors are generally used to make adjustments whose accuracy is unknown. The uncertainty factors comprise elements of the adjustment for uncertainty and variability. The default factors should be replaced with distributions that separate the elements transparently. Default distributions that characterize PK and PD variability, cross-species dose adjustments, and adjustments for the lack of sensitive studies will be needed as starting points that can be improved as the research advances.
- The committee considers that the term *safety factor*, to characterize uncertainty factors in noncancer risk assessments, is inappropriate and misleading. The term *uncertainty factor* is also inappropriate as it does not reflect the variability and adjustment elements that the factor represents.
- The underlying scientific and risk-management considerations point to the need for unification of cancer and noncancer approaches in which chemicals are put into a common analytic framework regardless of type of outcome. There are core differences among end points, but in this analytic framework a dose corresponding to a specified increase in risk in the population could be derived for both cancer and noncancer end points, and this would

Copyright © National Academy of Sciences. All rights reserved.

Science and Decisions: Advancing Risk Assessment
http://www.nap.edu/catalog/12209.html

add transparency and quantitative insight to risk-management decisions. Among other changes, this would involve a redefinition of the RfD. The committee acknowledges that the risk estimates and risk specific RfDs derived from this methodology will often be uncertain. This would nonetheless be an improvement over the RfDs derived from the traditional BMD and uncertainty factor approach. The results are more transparent, presenting variability and uncertainty, and are more amenable to refinements as better data are obtained. Further, quantification of risk (along with the attendant uncertainty) not only at the RfD but along the dose continuum is an important advance for risk benefit analysis.

- The committee finds that a common analytic framework best reflects the underlying science. The main elements of this framework are shown in Figure 5-8 and include

– Systematic assessments of the MOAs, vulnerable populations, and background exposures and disease processes that may affect a chemical's human dose-response relationships and human vulnerability. This includes an evaluation of the potential background exposures and processes (for example, damage and repair processes, disease, and aging) that interact with a chemical's MOAs and thus contribute to variability in and vulnerability to the toxicant response and that can result in a population dose-response relationship that is linear at low doses.

– Selection of a conceptual model for individual and population dose-response relationships. The following three are described in the chapter:

i. Low-dose nonlinear individual response, low-dose linear population response with background dependence.

ii. Low-dose nonlinear individual and population response independent of background.

iii. Low-dose linear individual and population response.

– Selection of a conceptual model and dose-response method that best reflects MOA and background considerations and the form of risk characterization needed for risk management. Where feasible, methods that result in quantitative descriptions of risk and uncertainty should be selected.

- The key advantages of the framework are

– Risk descriptors that are quantitative and probabilistic. The RfD would be redefined as a risk-specific dose (for example, the dose associated with a 1 in 100,000 risk of a particular end point), and the risk could be estimated at doses above and below the RfD. This would allow all end points to be more formally incorporated into risk-tradeoffs and benefit-cost analyses.

– Characterization of variability and uncertainty for critical end points. This would address concerns about population heterogeneity in risk and inform value-of-information and other priority-setting analyses that require quantitative uncertainty estimates. The sources of variability and uncertainty and their quantitative contributions in the derivation of risk estimates would be more transparent. This would in turn enable the quantitative characterization of uncertainties in such benefits.

– A means to quantitatively describe health benefits from changes in exposure. This would enable the direct comparisons of costs of these changes with the benefits accruing from them.

– The basis for more flexibility in decision-making. The risk manager can use the risk specific RfD in the same manner the current RfD is used in regulatory decision-making. However, additional quantitative risk information can accompany the RfD, including risk and uncertainty estimates above and below the RfD. This will enable a more robust consideration of options and trade-offs in risk-risk and risk-benefit analyses.

Copyright © National Academy of Sciences. All rights reserved.

Science and Decisions: Advancing Risk Assessment
http://www.nap.edu/catalog/12209.html

- The key disadvantages of the framework are
  - The need for increased analysis to consider in detail the background factors that may add to the exposure in question and that may contribute to variability. This can increase the complexity of the analysis and pose a challenge for communicating the analysis and its results. Training will be needed for both risk assessors and risk managers. The agency has already included some elements of the framework in a few assessments (for example, EPA 2001b; EPA 2004), and explored other elements in case studies (for example, Axelrad et al. 2005; Woodruff et al. 2007). EPA laboratories also conduct research that is supportive of the characterizations envisioned by the committee. Thus, EPA has internal capacity for the development of these methods. Realizing full use will take further development and staff training. The risk assessment community external to the agency provides several examples that are cited above and is also a resource for developing further cases and expanding the methodology. The agency also has considerable expertise translating risk information and using it in decision-making. Approaches currently used in risk management may have to be adapted to make full use of the new information and risk managers may need to be trained on how to best use the new and different risk characterizations.
  - Because of the limitations of data on which some elements of the framework would be built, this necessarily entails development of defaults. Depending on the level of analysis, that would provide incentives for chemical-specific information on background exposures, interaction with baseline aging and disease processes, and interindividual variability. It comes at a time when toxicology and risk-assessment resources are already challenged by the expanding role of risk assessment in decision-making and the lack of basic toxicology information on many chemicals. However, it also comes at a time of rapid scientific and technologic innovation in the biologic sciences and testing that can be developed to support novel and improved approaches (NRC 2006d, 2007a,b).
- Establishing reasonable and scientifically supported default approaches (such as linear extrapolation to low dose for chemicals that are subject to background additivity) and default distributions (such as interindividual variability) to implement the framework will encourage research and a healthy discussion of the science that underpins risk assessment. The resulting default approaches are part of the anticipated advances in the use of defaults in risk assessment described in Chapter 6. The process of establishing the defaults will bring about a better understanding of how chemical-specific information should be used to inform toxicity assessment and low-dose extrapolation.

Recommendations

The committee has divided its recommendations on the unified framework into short- and long- term recommendations. If the short term recommendations are implemented, the committee envisions significant progress in the next 2-5 years. The time horizon for substantial progress for the long term recommendations is further out, 10-20 years.

Short-Term Recommendations

- The committee recommends the phase-in of the unified framework for dose-response assessment as new chemicals are assessed or old ones are reassessed for Integrated Risk Information System or program offices or incorporated in comparative or cost-benefit analyses. The initial test cases should be used as a proof of concept. The committee recommends a flexible approach in which different conceptual models can be applied in the unified frame-

Copyright © National Academy of Sciences. All rights reserved.

Science and Decisions: Advancing Risk Assessment
http://www.nap.edu/catalog/12209.html

work, as illustrated by the three conceptual models presented in this chapter. This approach would involve

– Incorporation of probabilistic and distributional methods into dose-response analysis for agents believed to have low-dose nonlinear responses and the later redefinition of the RfD on the basis of the probabilistic description.

– Evaluation of each chemical in terms of MOA, background exposure and disease processes, and vulnerable populations. This would add a step to the dose-response analysis in which background exposures and vulnerabilities of the target population are analyzed and used to decide between analytic options based on conceptual models, according to the unified framework outlined in Figure 5-8.

– Incorporation of background additivity to account for

o Additional sources of exposure to the same chemical or to similarly acting chemicals (including endogenous sources).

o Chemical MOA interaction with relevant disease or aging processes that lead to a background vulnerability distribution.

– Development of defaults and guidance for assessing the MOA, background exposure and disease processes, and vulnerable populations, and selection of conceptual model. The committee recommends that cancer and noncancer responses be assumed to be linear as a default. An alternative analytic option (conceptual model 2) is available for cases in which it can be shown that background is unlikely to be an important contributor to risk, according to the recommended evaluation of MOAs and background.

– Formal introduction of human variability into cancer dose-response modeling and risk characterization. This will require chemical-specific distributions or the use of default variability distributions. The committee recommends that as the distributions are being developed, EPA use a default for interindividual variability that assumes a lognormal distribution and immediately begin to explicitly address human variability in cancer response estimates. A reasonable assumption would be that the 95% upper-bound person is about 10-50 times as sensitive as the median person.

- The committee recommends that EPA develop case studies to explore the use of the new unified framework. The goal of the case studies would not be simply to compare the results of the current approach and new framework. Rather, the case studies would be used to explore and gain experience with the framework in the MOA, vulnerability, and background assessments; using improved information on variability (for example, genetic polymorphisms, disease, and aging-related vulnerabilities) and coexposures in RfD derivation; incorporating variability into cancer risk analysis; and quantitative uncertainty characterizations of dose-response relationships.
- The committee recommends that EPA gather information from epidemiology, the pharmaceutical literature, and clinical toxicology and use it to develop default interhuman variability PK and PD distributions. Some possible approaches are outlined in Table 5-1.
- The committee recommends that the agency develop default-adjustment distributions that quantitatively characterize the adjustments and key uncertainties typical in dose-response assessment, including cross-species extrapolation in PK and PD and extrapolations among dose route, dosing intervals (for example, subchronic to chronic), and data gaps. Some possible approaches are outlined in Table 5-1. Maximum use of existing human datasets is encouraged. Studies with well-defined exposure information, such as biomarker measurements on individuals, could be examined to understand the heterogeneity in response. Such datasets could be used to build variability distributions that may be applicable to sets of chemicals (with similar structure, MOA, target sites, and effects) and increase understanding of interhuman PD variability.

Copyright © National Academy of Sciences. All rights reserved.

Science and Decisions: Advancing Risk Assessment
http://www.nap.edu/catalog/12209.html

- The agency should develop formal guidance for dose-response analysis under the unified framework. For example, guidance will be needed for the conduct of background vulnerability and exposure assessments, MOA evaluation, default dose-response modeling, nondefault chemical specific analyses.
- The committee recommends as default distributions are developed for the different adjustments used in dose-response assessment, they should be assigned accurate labels (such as *human variability distribution*). This should lessen the opportunity for transferring to the new default distributions the misunderstanding commonly associated with use of the term *uncertainty factor*.
- Over the next 5 years, the committee recommends that EPA further develop the issue of vulnerability by gathering data and developing a broad array of human-vulnerability information from the biomedical literature, focusing on diseases that are likely to interact with the MOAs of prevalent-exposure and high-priority chemicals (for example, pulmonary, cardiovascular, hepatic, and renal diseases and various cancers). This could involve working with clinicians, biochemists, epidemiologists, and other biomedical specialists to develop preclinical-disease biomarkers as upstream indicators of vulnerability to toxicant MOAs.

Long-Term Recommendations

- The committee recommends that EPA expand its research on the issues of vulnerability and susceptibility. The agency could conduct studies itself and coordinate with other agencies for more in-depth research on the determinants of vulnerability and the development of approaches for more accurate consideration of vulnerability in agency assessments. This could involve using epidemiologic studies to explore how the response to toxicants may be affected by pre-existing diseases and vulnerabilities in the population. Biomarkers of vulnerability and effect could be developed for applications as predictive screens in exposed populations. When analyzed with exposure biomarkers, they could be used to assess human exposure-response relationships and interindividual variability. Regional and national datasets, such as those from National Health and Nutrition Examination Surveys and environmental and public-health tracking, could be used to evaluate whether people with background vulnerability or background exposure are at increased risk of the effects of exposure to toxicants. This work could lead to vulnerability distributions for use in dose-response assessment. Pharmacogenetic and polymorphism probes could be incorporated into epidemiologic studies to explore key interindividual susceptibility factors and their frequency in the population. Animal models, such as genetically modified knockout mice, could be used to define the functional importance of particular genes and their polymorphisms in determining risk.
- The committee recommends computational research that applies systems-biology techniques to analyze how -omics end points might inform the development of distributions outlined in Table 5-1. For example, analyzing data from high-throughput screens with genomics end points may result in interpretable upstream indicators of disease vulnerability. The biochemical processes that lead to pathologic conditions or functional loss could be described by continuous parameters that may be suitable as disease biomarkers in the population. These approaches could also provide interpretable biochemical end points reflective of key steps in a toxicant's MOA.
- The committee recommends exploration into interactions of exposures to chemicals that have similar or different MOAs but affect the same toxicologic process. Such research should improve understanding of issues related to background additivity. The research would also affect approaches to mixtures and combined exposures and to the question of whether it

Copyright © National Academy of Sciences. All rights reserved.

Science and Decisions: Advancing Risk Assessment
http://www.nap.edu/catalog/12209.html

is more appropriate to assume effect additivity (now assumed in noncancer risk assessment), dose additivity, or some other characteristic in a given risk-assessment circumstance.

REFERENCES

Angelosanto, F.A., G.R. Blackburn, C.A. Schreiner, and C.R. Mackerer. 1996. Benzene induces a dose-responsive increase in the frequency of micronucleated cells in rat Zymbal glands. Environ. Health Perspect. 104(Suppl. 6):1331-1336.

Autrup, H., B. Daneshvar, L.O. Dragsted, M. Gamborg, A.M. Hansen, S. Loft, H. Okkels, F. Nielsen, P.S. Nielsen, E. Raffn, H. Wallin, and L.E. Knudsen. 1999. Biomarkers for exposure to ambient air pollution—comparison of carcinogen-DNA adduct levels with other exposure markers and markers for oxidative stress. Environ. Health Perspect. 107(3):233-238.

Axelrad, D.A., K. Baetcke, C. Dockins, C.W. Griffiths, R.N. Hill, P.A. Murphy, N. Owens, N.B. Simon, and L.K. Teuschler. 2005. Risk assessment for benefits analysis: Framework for analysis of a thyroid-disrupting chemical. J. Toxicol. Environ. Health A. 68(11-12):837-855.

Axelrad, D.A., D.C. Bellinger, L.M. Ryan, and T.J. Woodruff. 2007. Dose-response relationship of prenatal mercury exposure and IQ: An integrative analysis of epidemiologic data. Environ. Health Perspect. 115(4):609-615.

Baird, S.J.S., J.T. Cohen, J.D. Graham, A.I. Shiyakhter, and J.S. Evans. 1996. Non-cancer risk assessment: A probabilistic alternative to current practice. Hum. Ecol. Risk Asses. 2(1):79-102.

Bannasch, P. 2003. Comments on 'R. Karbe and R. L. Kerlin. (2002). Cystic Degeneration/Spongiosis Hepatis [Toxicol. Pathol. 30(2):216-227].' Toxicol. Pathol. 31(5):566-570.

Bois, F.Y., G. Krowech, and L. Zeise. 1995. Modeling human interindividual variability in metabolism and risk: The example of 4-aminobiphenyl. Risk Anal. 15(2):205-213.

Boutet, K, J.L. Malo, H. Ghezzo, and D. Gautrin. 2007. Airway hyperresponsiveness and risk of chest symptoms in an occupational model. Thorax 62(3):260-264.

Brouwer, A., M.P. Longnecker, L.S. Birnbaum, J. Cogliano, P. Kostyniak, J. Moore, S. Schantz, and G. Winneke. 1999. Characterization of potential endocrine-related health effects at low-dose levels of exposure to PCBs. Environ. Health Perspect. 107(Suppl. 4):639-649.

CDC (Centers for Disease Control and Prevention). 2007. Asthma: Asthma's Impact on Children and Adolescents. U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, Washington, DC [online]. Available: http://www.cdc.gov/asthma/children.htm [accessed Sept. 14, 2007].

Crofton, K.M., E.S. Craft, J.M. Hedge, C. Gennings, J.E. Simmons, R.A. Carchman, W.H. Carter Jr., and M.J. DeVito. 2005. Thyroid-hormone-disrupting chemicals: Evidence for dose-dependent additivity or synergism. Environ. Health Perspect. 113(11):1549-1554.

Crouch, E., and R. Wilson. 1979. Interspecies comparison of carcinogenic potency. J. Toxicol. Environ. Health 5(6):1095-1118.

Crump, K.S., D.G. Hoel, C.H. Langley, and R. Peto. 1976. Fundamental carcinogenic processes and their implications for low dose risk assessment. Cancer Res. 36(9 Pt. 1):2973-2979.

Daniels, M.J., F. Dominici, J.M. Samet, and S.L. Zeger. 2000. Estimating particulate matter-mortality dose-response curves and threshold levels: An analysis of daily time-series for the 20 largest U.S. cities. Am. J. Epidemiol. 152(5):397-406.

de Meer, G., D.S. Postma, and D. Heederik. 2003. Bronchial responsiveness to adenosine-5'-monophosphate and methacholine as predictors for nasal symptoms due to newly introduced allergens: A follow-up study among laboratory animal workers and bakery apprentices. Clin. Exp. Allergy 33(6):789-794.

Demchuk, E., B. Yucesoy, V.J. Johnson, M. Andrew, A. Weston, D.R. Germolec, C.T. De Rosa, and M.I. Luster. 2007. A statistical model for assessing genetic susceptibility as a risk factor in multifactorial diseases: Lessons from occupational asthma. Environ. Health Perspect. 115(2):231-234.

Dominici, F., A. McDermott, S.L. Zeger, and J.M. Samet. 2003. National maps of the effects of particulate matter on mortality: Exploring geographical variation. Environ. Health Perspect. 111(1):39-43.

Dubowsky, S.D., H. Suh, J. Schwartz, B.A. Coull, and D.R. Gold. 2006. Diabetes, obesity, and hypertension may enhance associations between air pollution and markers of systemic inflammation. Environ. Health Perspect. 114(7):992-998.

EPA (U.S. Environmental Protection Agency). 1993. Alachlor (CASRN 15972-60-8). Integrated Risk Information System, U.S. Environmental Protection Agency, Washington, DC [online]. Available: http://www.epa.gov/iris/subst/0129.htm [accessed Sept. 12, 2007].

Copyright © National Academy of Sciences. All rights reserved.

Science and Decisions: Advancing Risk Assessment
http://www.nap.edu/catalog/12209.html

EPA (U.S. Environmental Protection Agency). 1998. Health Risks from Low-Level Environmental Exposure to Radionuclides. Federal Guidance Report No.13 -Part 1, Interim Version. EPA 402-R-97-014. Office of Radiation and Indoor Air, U.S. Environmental Protection Agency, Washington, DC. January 1998 [online]. Available: http://homer.ornl.gov/VLAB/fedguide13.pdf [accessed Aug. 7, 2008].

EPA (U.S. Environmental Protection Agency). 2000a. Benchmark Dose Technical Guidance Document. EPA/630/R-00/001. Risk Assessment Forum, U.S. Environmental Protection Agency, Washington, DC. October 2000 [online]. Available: http://www.epa.gov/ncea/pdfs/bmds/BMD-External_10_13_2000.pdf [accessed Jan. 4, 2008].

EPA (U.S. Environmental Protection Agency). 2000b. Risk Characterization Handbook. EPA 100-B-00-002. Office of Science Policy, Office of Research and Development, U.S. Environmental Protection Agency, Washington, DC. December 2000 [online]. Available: http://www.epa.gov/OSA/spc/pdfs/rchandbk.pdf [accessed Jan. 7, 2008].

EPA (U.S. Environmental Protection Agency). 2001a. Trichloroethylene Health Risk Assessment: Synthesis and Characterization. External Review Draft. EPA/600/P-01/002A. Office of Research and Development, Washington, DC. August 2001 [online]. Available: http://rais.ornl.gov/tox/TCEAUG2001.PDF [accessed Aug. 2, 2008].

EPA (U.S. Environmental Protection Agency). 2001b. Water Quality Criterion for the Protection of Human Health: Methylmercury, Chapter 4: Risk Assessment for Methylmercury. EPA-823-R-01-001. Office of Science and Technology, Office of Water, U.S. Environmental Protection Agency, Washington, DC. January 2001 [online]. Available: http://www.epa.gov/waterscience/criteria/methylmercury/pdf/merc45.pdf [accessed Aug. 6, 2008].

EPA (U.S. Environmental Protection Agency). 2001c. Chloroform (CASRN 67-66-3). Integrated Risk Information System, U.S. Environmental Protection Agency, Washington, DC [online]. Available: http://www.epa.gov/iris/subst/0025.htm [accessed Sept. 12, 2007].

EPA (U.S. Environmental Protection Agency). 2002a. A Review of the Reference Dose and Reference Concentration Processes. EPA/630/P-02/002F. Risk Assessment Forum, U.S. Environmental Protection Agency, Washington, DC. December 2002 [online]. Available: http://cfpub.epa.gov/ncea/cfm/recordisplay.cfm?deid=55365 [accessed Jan. 4, 2008].

EPA (U.S. Environmental Protection Agency). 2002b. Guidance on Cumulative Risk Assessment of Pesticide Chemicals That Have a Common Mechanism of Toxicity. Office of Pesticide Programs, U.S. Environmental Protection Agency, Washington, DC. January 14, 2002 [online]. Available: http://www.epa.gov/pesticides/trac/science/cumulative_guidance.pdf [accessed Jan. 7, 2008].

EPA (U.S. Environmental Protection Agency). 2004. Exposure and Human Health Reassessment of 2,3,7,8-Tetrachlorodibenzo-p-Dioxin (TCDD) and Related Compounds, NAS Review Draft. EPA/600/P-00/001Cb. National Center for Environmental Assessment, Office of Research and Development, U.S. Environmental Protection Agency, Washington, DC [online]. Available: http://cfpub.epa.gov/ncea/cfm/recordisplay.cfm?deid=87843 [accessed Aug. 6, 2008].

EPA (U.S. Environmental Protection Agency). 2005a. Approaches for the Application of Physiologically Based Pharmacokinetic (PBPK) Models and Supporting Data in Risk Assessment. External Review Draft. EPA/600/R-05/043A. National Center for Environmental Assessment, Office of Research and Development, U.S. Environmental Protection Agency, Washington, DC. June 2005 [online]. Available: http://cfpub.epa.gov/ncea/cfm/recordisplay.cfm?deid=135427 [accessed Jan. 22, 2008].

EPA (U.S. Environmental Protection Agency). 2005b. Guidelines for Carcinogen Risk Assessment. EPA/630/P-03/001F. Risk Assessment Forum, U.S. Environmental Protection Agency, Washington, DC. March 2005 [online]. Available: http://cfpub.epa.gov/ncea/cfm/recordisplay.cfm?deid=116283 [accessed Feb. 7, 2007].

EPA (U.S. Environmental Protection Agency). 2005c. Supplemental Guidance for Assessing Susceptibility for Early-Life Exposures to Carcinogens. EPA/630/R-03/003F. Risk Assessment Forum, U.S. Environmental Protection Agency, Washington, DC. March 2005 [online]. Available: http://cfpub.epa.gov/ncea/cfm/recordisplay.cfm?deid=160003 [accessed Jan. 4, 2008].

EPA (U.S. Environmental Protection Agency). 2005d. Toxicological Review of Phosgene (CAS 75-44-5) In Support of Summary Information on the Integrated Risk Information System (IRIS). EPA/635/R-06/001. U.S. Environmental Protection Agency, Washington, DC. December 2005 [online]. Available: http://www.epa.gov/iris/toxreviews/0487-tr.pdf [accessed Aug. 7, 2008].

EPA (U.S. Environmental Protection Agency). 2006a. Background on Risk Characterization. Technology Transfer Network, 1999 National-Scale Air Toxics Assessment, U.S. Environmental Protection Agency, Washington, DC [online]. Available: http://www.epa.gov/ttn/atw/nata1999/riskbg.html [accessed Sept. 10, 2007].

EPA (U.S. Environmental Protection Agency). 2006b. Harmonization in Interspecies Extrapolation: Use of BW¾ as Default Method in Derivation of the Oral RfD. External Review Draft. EPA/630/R-06/001. Risk Assessment Forum Technical Panel, U.S. Environmental Protection Agency, Washington, DC. February 2006 [online]. Available: http://cfpub.epa.gov/si/si_public_record_Report.cfm?dirEntryID=148525 [accessed Aug. 7, 2008].

Copyright © National Academy of Sciences. All rights reserved.

Science and Decisions: Advancing Risk Assessment
http://www.nap.edu/catalog/12209.html

EPA (U.S. Environmental Protection Agency). 2006c. Toxicological Review of Dibutyl Phthalate (Di-n-Phthalate) (CAS No. 84-74-2) In Support of Summary Information on the Integrated Risk Information System (IRIS). NCEA-S-1755. U.S. Environmental Protection Agency, Washington, DC [online]. Available: http://cfpub.epa.gov/ncea/cfm/recordisplay.cfm?deid=155707 [accessed Dec. 10, 2007].

EPA SAB (U.S. Environmental Protection Agency Science Advisory Board). 2000. Review of the EPA'S Draft Chloroform Risk Assessment. EPA-SAB-EC-00-009. Science Advisory Board, U.S. Environmental Protection Agency, Washington, DC. April 2000 [online]. Available: http://yosemite.epa.gov/sab/sabproduct.nsf/D0E41CF58569B1618525719B0064BC3A/$File/ec0009.pdf [accessed Jan. 22, 2008].

EPA SAB (U.S. Environmental Protection Agency Science Advisory Board). 2002. Workshop on the Benefits of Reductions in Exposure to Hazardous Air Pollutants: Developing Best Estimates of Dose-Response Functions. EPA-SAB-EC-WKSHP-02-001. Science Advisory Board, U.S. Environmental Protection Agency, Washington, DC. January 2002 [online]. Available: http://yosemite.epa.gov/sab/sabproduct.nsf/34355712EC011A358525719A005BF6F6/$File/ecwkshp02001+appa-g.pdf [accessed Jan. 22, 2008].

Evans, J.S., and S.J.S. Baird. 1998. Accounting for missing data in noncancer risk assessment. Hum. Ecol. Risk Assess. 4(2):291-317.

Evans, J.S., L.R. Rhomberg, P.L. Williams, A.M. Wilson, and S.J.S. Baird. 2001. Reproductive and developmental risks from ethylene oxide: A probabilistic characterization of possible regulatory thresholds. Risk Anal. 21(4):697-717.

Freireich, E.J., E.A. Gehan, D.P. Rall, L.H. Schmidt, and H.E. Skipper. 1966. Quantitative comparison of toxicity of anti-cancer agents in mouse, rat, hamster, dog, monkey, and man. Cancer Chemother. Rep. 50(4):219-244.

Gaylor, D., L. Ryan, D. Krewski, and Y. Zhu. 1998. Procedures for calculating benchmark doses for health risk assessment. Regul. Toxicol. Pharmacol. 28(2):150-164.

Gaylor, D.W., R.L. Kodell, J.J. Chen, and D. Krewski. 1999. A unified approach to risk assessment for cancer and noncancer endpoints based on benchmark doses and uncertainty/safety factors. Regul. Toxicol. Pharmacol. 29(2 Pt 1):151-157.

Ginsberg, G., D. Hattis, B. Sonawane, A. Russ, P. Banati, M. Kozlak, S. Smolenski, and R. Goble. 2002. Evaluation of child/adult pharmacokinetic differences from a database derived from the therapeutic drug literature. Toxicol. Sci. 66(2):185-200.

Gray, L.E., J. Ostby, J. Furr, C.J. Wolf, C. Lambright, L. Parks, D.N. Veeramachaneni, V. Wilson, M. Price, A. Hotchkiss, E. Orlando, and L. Guillette. 2001. Effects of environmental antiandrogens on reproductive development in experimental animals. Hum. Reprod. Update 7(3):248-264.

Gu, J., D. Liang, Y. Wang, C. Lu, X. Wu. 2005. Effects of N-acetyl transferase 1 and 2 polymorphisms on bladder cancer risk in Caucasians. Mutat Res. 581(1-2):97-104.

Hass, U., M. Scholze, S. Christiansen, M. Dalgaard, A.M. Vinggaard, M. Axelstad, S.B. Metzdorff, and A. Kortenkamp. 2007. Combined exposure to anti-androgens exacerbates disruption of sexual differentiation in the rat. Environ. Health Perspect. 115(Suupl.1):122-128.

Hattis, D. 2008. Distributional analyses for children's inhalation risk assessments. J. Toxicol. Environ. Health Part A 71(3):218-226.

Hattis, D., and K. Barlow. 1996. Human interindividual variability in cancer risks -technical and management challenges. Hum. Ecol. Risk Assess. 2(1):194-220.

Hattis, D., and M.K. Lynch. 2007. Empirically observed distributions of pharmacokinetic and pharmacodynamic variability in humans: Implications for the derivation of single point component uncertainty factors providing equivalent protection as existing RfDs. Pp. 69-93 in Toxicokinetics and Risk Assessment, J.C. Lipscomb, and E.V. Ohanian, eds. New York: Informa Healthcare.

Hattis, D., A. Russ, R. Goble, P. Banati, and M. Chu. 2001. Human interindividual variability in susceptibility to airborne particles. Risk Anal. 21(4):585-599.

Hattis, D., S. Baird, and R. Goble. 2002. A straw man proposal for a quantitative definition of the RfD. Drug Chem. Toxicol. 25(4):403-436.

Hattis, D., G. Ginsberg, B. Sonawane, S. Smolenski, A. Russ, M. Kozlak, and R. Goble. 2003. Differences in pharmacokinetics between children and adults. II. Children's variability in drug elimination half-lives and in some parameters needed for physiologically-based pharmacokinetic modeling. Risk Anal. 23(1):117-142.

Hsiao, T.J., J.D. Wang, P.M. Yang, P.C. Yang, and T.J. Cheng. 2004. Liver fibrosis in asymptomatic polyvinyl chloride workers. J. Occup. Environ. Med. 46(9):962-966.

IARC (International Agency for Research on Cancer). 2000. Di(2-ethylhexyl)phthalate. Pp. 41-148 in Some Industrial Chemicals. IARC Monographs on the Evaluation of Carcinogenic Risks to Humans Vol. 77. Lyon: IARC.

Copyright © National Academy of Sciences. All rights reserved.

Science and Decisions: Advancing Risk Assessment
http://www.nap.edu/catalog/12209.html

IARC (International Agency for Research on Cancer). 2006. Preamble. IARC Monographs on the Evaluation of Carcinogenic Risks to Humans. International Agency for Research on Cancer, World Health Organization, Lyon [online]. Available: http://monographs.iarc.fr/ENG/Preamble/CurrentPreamble.pdf [accessed Aug. 6, 2008].

ICRP (International Commission on Radiological Protection). 1991. 1990 Recommendations of the International Commission on Radiological Protection. ICRP Publication 60. Annals of the ICPR 21(1/3). New York: Pergamon.

ICRP (International Commission on Radiological Protection). 2006. Low Dose Extrapolation of Radiation-Related Cancer Risk. ICRP Publication 99. Annals of the ICPR 35(4). Oxford: Elsevier.

Inatomi, H., T. Katoh, T. Kawamoto, and T. Matsumoto. 1999. NAT2 gene polymorphism as a possible marker for susceptibility to bladder cancer in Japanese. Int. J. Urol. 6(9):446-454.

IPCS (International Programme for Chemical Safety). 2005. Guidance for the use of data in development of chemical-specific adjustment factors for interspecies differences and human variability. Pp. 25-48 in Chemical-Specific Adjustment Factors for Interspecies Differences and Human Variability: Guidance Document for Use of Data in Dose/Concentration-Response Assessment. Harmonization Project Document No. 2. International Programme for Chemical Safety, World Health Organization, Geneva [online]. Available: http://whqlibdoc.who.int/publications/2005/9241546786_eng.pdf [accessed Aug. 6, 2008].

Ito, Y., O. Yamanoshita, N. Asaeda, Y. Tagawa, C.H. Lee, T. Aoyama, G. Ichihara, K. Furuhashi, M. Kamijima, F.J. Gonzalez, and T. Nakajima. 2007. Di(2-ethylhexyl)phthalate induces hepatic tumorigenesis through a peroxisome proliferator-activated receptor alpha-independent pathway. J. Occup. Health. 49(3):172-182.

Jeong, Y.C., N.J. Walker, D.E. Burgin, G. Kissling, M. Gupta, L. Kupper, L.S. Birnbaum, and J.A. Swenberg. 2008. Accumulation of M(1)dG DNA adducts after chronic exposure to PCBs, but not from acute exposure to polychlorinated aromatic hydrocarbons. Free Radic Biol. Med. 45(5):585-591.

Jiang, X, J.M. Yuan, P.L. Skipper, S.R. Tannenbaum, and M.C. Yu. 2007. Environmental tobacco smoke and bladder cancer risk in never smokers of Los Angeles county. Cancer Res. 67(15):7540-7545.

Kang, D.H., N. Rothman, M.C. Poirier, A. Greenberg, C.H. Hsu, B.S. Schwartz, M.E. Baser, J.D. Groopman, A. Weston, and P.T. Strickland. 1995. Interindividual differences in the concentration of 1-hydroxypyrene-glucuronide in urine and polycyclic aromatic hydrocarbon-DNA adducts in peripheral white blood cells after charbroiled beef consumption. Carcinogenesis 16(5):1079-1085.

Karbe, R., and R.L. Kerlin. 2002. Cystic degeneration/Spongiosis hepatis in rats. Toxicol. Pathol. 30(2):216-227.

Klaunig, J.E., M.A. Babich, K.P. Baetcke, J.C. Cook, J.C. Corton, R.M. David, J.G. DeLuca, D.Y. Lai, R.H. McKee, J.M. Peters, R.A. Roberts, and P.A. Fenner-Crisp. 2003. PPARalpha agonist-induced rodent tumors: Modes of action and human relevance. Crit. Rev. Toxicol. 33(6):655-780.

Kobayashi, E., Y. Suwazono, M. Uetani, T. Inaba, M. Oishi, T. Kido, M. Nishijo, H. Nakagawa, and K. Nogawa. 2006. Estimation of benchmark dose as the threshold levels of urinary cadmium, based on excretion of total protein, beta2-microglobulin, and N-acetyl-beta-D-glucosaminidase in cadmium nonpolluted regions in Japan. Environ. Res. 101(3):401-406.

Kodavanti, U.P., D.L. Costa, S.N. Giri, B. Starcher, and G.E. Hatch. 1997. Pulmonary structural and extracellular matrix alterations in Fischer F344 rats following subchronic phosgene exposure. Fundam. Appl. Toxicol. 37(1):54-63.

Kramer, C.B., A.C. Cullen, and E.M. Faustman. 2006. Policy implications of genetic information on regulation under the Clean Air Act: The case of particulate matter and asthmatics. Environ. Health Perspect. 114(3):313-319.

Laprise, C., and L.P. Boulet. 1997. Asymptomatic airway hyperresponsiveness: A three-year follow-up. Am. J. Respir. Crit. Care Med. 156(2 Pt. 1):403-409.

Lutz, W.K. 1990. Dose-reponse relationship and low dose extrapolation in chemical carcinogenesis. Carcinogenesis 11(8):1243-1247.

Lutz, W.K. 2001. Susceptibility differences in chemical carcinogenesis linearize the dose-response relationship: Threshold doses can be defined only for individuals. Mutat. Res. 482(1-2):71-76.

Maroni, M., and A.C. Fanetti. 2006. Liver function assessment in workers exposed to vinyl chloride. Int. Arch. Occup. Environ. Health 79(1):57-65.

Metzdorff, S.B., M. Dalgaard, S. Christiansen, M. Axelstad, U. Hass, M.K. Kiersgaard, M. Scholze, A. Kortenkamp, and A.M. Vinggaard. 2007. Dysgenesis and histological changes of genitals and perturbations of gene expression in male rats after in utero exposure to antiandrogen mixtures. Toxicol. Sci. 98(1):87-98.

Nakajima, M., M. Itoh, H. Sakai, T. Fukami, M. Katoh, H. Yamazaki, F.F. Kadlubar, S. Imaoka, Y. Funae, and T. Yokoi. 2006. CYP2A13 expressed in human bladder metabolically activates 4-aminobiphenyl. Int. J. Cancer 119(11):2520-2526.

Nessel, C.S., S.C. Lewis, K.L. Stauber, and J.L. Adgate. 1995. Subchronic to chronic exposure extrapolation: Toxicologic evidence for a reduced uncertainty factor. Hum. Ecol. Risk Assess. 1(5):516-526.

Copyright © National Academy of Sciences. All rights reserved.

Science and Decisions: Advancing Risk Assessment
http://www.nap.edu/catalog/12209.html

NRC (National Research Council). 1993. Pesticides in the Diets of Infants and Children. Washington, DC: National Academy Press.

NRC (National Research Council). 1994. Science and Judgment in Risk Assessment. Washington, DC: National Academy Press.

NRC (National Research Council). 1998a. Assessment of Exposure-Response Functions for Rocket-Emission Toxicants. Washington, DC: National Academy Press.

NRC (National Research Council). 1998b. Health Effects of Exposure to Low Levels of Ionizing Radiations: Time for Reassessment? Washington, DC: National Academy Press.

NRC (National Research Council). 2000. Copper in Drinking Water. Washington, DC: National Academy Press.

NRC (National Research Council). 2006a. Assessing the Human Risks of Trichloroethylene. Washington, DC: The National Academies Press.

NRC (National Research Council). 2006b. Health Risks for Dioxin and Related Compounds. Washington, DC: The National Academies Press.

NRC (National Research Council). 2006c. Health Risks from Exposures to Low Levels of Ionizing Radiation: BEIR VII. Washington, DC: The National Academies Press.

NRC (National Research Council). 2006d. Toxicity Testing for Assessment of Environmental Agents: Interim Report. Washington, DC: The National Academies Press.

NRC (National Research Council). 2007a. Toxicity Testing in the Twenty-First Century: A Vision and a Strategy. Washington, DC: The National Academies Press.

NRC (National Research Council). 2007b. Applications of Toxicogenomic Technologies to Predictive Toxicology and Risk Assessment. Washington, DC: The National Academies Press.

Peel, J.L, K.B. Metzger, M. Klein, W.D. Flanders, J.A. Mulholland, and P.E. Tolbert. 2007. Ambient air pollution and cardiovascular emergency department visits in potentially sensitive groups. Am. J. Epidemiol. 165(6):625-633.

Rhomberg, L.R., and S.K. Wolff. 1998. Empirical scaling of single oral lethal doses across mammalian species based on a large database. Risk Anal. 18(6):741-753.

Sackey, K. 1999. Hemolytic anemia: Part 1. Pediatr. Rev. 20(5):152-158.

Samoli, E., A. Analitis, G. Touloumi, J. Schwartz, H.R. Anderson, J. Sunyer, L. Bisanti, D. Zmirou, J.M. Vonk, J. Pekkanen, P. Goodman, A. Paldy, C. Schindler, and K. Katsouyanni. 2005. Estimating the exposure–response relationships between particulate matter and mortality within the APHEA multicity project. Environ. Health Perspect. 113(1):88-95.

Sand, S.J., D. von Rosen, and A.F. Filipsson. 2003. Benchmark dose calculations in risk assessment using continuous dose-response information: The influence of variance and the determinants of a cut-off value. Risk Anal. 23(5):1059-1068.

Santella, R.M., K. Hemminki, D.L. Tang, M. Paik, R. Ottman, T.L Young, K. Savela, L. Vodickova, C. Dickey, R. Whyatt, and F.P. Perera. 1993. Polycyclic aromatic hydrocarbon-DNA adducts in white blood cells and urinary 1-hydroxypyrene in foundry workers. Cancer Epidemiol. Biomarkers Prev. 2(1):59-62.

Schwartz, J., and A. Zanobetti. 2000. Using meta-smoothing to estimate dose-response trends across multiple studies, with application to air pollution and daily death. Epidemiology 11(6):666-672.

Schwartz, J., F. Laden, and A. Zanobetti. 2002. The concentration–response relation between $PM_{2.5}$ and daily deaths. Environ. Health Perspect. 110(10):1025-1029.

Schwartz, J., B. Coull, F. Laden, and L. Ryan. 2008. The effect of dose and timing of dose on the association between airborne particles and survival. Environ. Health Perspect. 116(1):64-69.

Siwinska, E., D. Mielzynska, and L. Kapka. 2004. Association between urinary 1-hydroxypyrene and genotoxic effects in coke oven workers. Occup. Environ. Med. 61(3):e10.

Slikker, W. Jr, M.E. Andersen, M.S. Bogdanffy, J.S. Bus, S.D. Cohen, R.B. Conolly, R.M. David, N.G. Doerrer, D.C. Dorman, D.W. Gaylor, D. Hattis, J.M. Rogers, R. Setzer, J.A. Swenberg, and K. Wallace. 2004. Dose-dependent transitions in mechanisms of toxicity. Toxicol. Appl. Pharmacol. 201(3):203-225.

Subramaniam, R.P., P. White, and V.J. Cogliano. 2006. Comparison of cancer slope factors using different statistical approaches. Risk Anal. 26(3):825-830.

Travis, C.C., and R.K. White. 1988. Interspecific scaling of toxicity data. Risk Anal. 8(1):119-125.

Tsuneoka, Y., T.P. Dalton, M.L. Miller, C.D. Clay, H.G. Shertzer, G. Talaska, M. Medvedovic, and D.W. Nebert. 2003. 4-aminobiphenyl–induced liver and urinary bladder DNA adduct formation in Cyp1a2(–/–) and Cyp1a2(+/+) mice. J. Natl. Cancer Inst. 95(16):1227-1237.

Watanabe, K., F.Y. Bois, and L. Zeise. 1992. Interspecies extrapolation: A reexamination of acute toxicity data. Risk Anal. 12(2):301-310.

Weil, C.S., and D.D. McCollister. 1963. Safety evaluation of Chemicals: Relationship between short- and long-term feeding studies in designing an effective toxicity test. Agr. Food Chem. 11(6):486-491.

Copyright © National Academy of Sciences. All rights reserved.

Science and Decisions: Advancing Risk Assessment
http://www.nap.edu/catalog/12209.html

West, J., H. Wood, R.F. Logan, M. Quinn, and G.P. Aithal. 2006. Trends in the incidence of primary liver and biliary tract cancers in England and Wales 1971-2001. Br. J. Cancer 94(11):1751-1758.

Wolf, C.J., G.A. LeBlanc, and L.E. Gray Jr. 2004. Interactive effects of vinclozolin and testosterone propionate on pregnancy and sexual differentiation of the male and female SD rat. Toxicol. Sci. 78(1):135-143.

Woodruff, T.J., E.M. Wells, E.W. Holt, D.E. Burgin, and D.A. Axelrad. 2007. Estimating risk from ambient concentrations of acrolein across the United States. Environ. Health Perspect. 115(3):410-415.

Wrixon, A.D. 2008. New ICRP recommendations. J. Radiol. Prot. 28(2):161-168.

Yamazaki, K, H. Ohno, M. Asakura, A. Narumi, H. Ohbayashi, H. Fujita, M. Ohnishi, T. Katagiri, H. Senoh, K. Yamanouchi, E. Nakayama, S. Yamamoto, T. Noguchi, K. Nagano M. Enomoto, and H. Sakabe. 1994. Two-year toxicological and carcinogenesis studies of 1, 4-dioxane in F344 rats and BDF1 mice: Drinking studies. Pp. 193-198 in Proceedings of the Second Asia-Pacific Symposium on Environmental and Occupational Health, 22-24 July, 1993, Kobe, Japan, K. Sumino, and S. Sato, eds. Kobe: International Center for Medical Research Kobe, University School of Medicine.

Zanobetti, A., J. Schwartz, and D. Gold. 2000. Are there sensitive subgroups for the effects of airborne particles? Environ. Health Perspect. 108(9):841-845.

Copyright © National Academy of Sciences. All rights reserved.