Human Health Risk Assessment

RONALD E. BAYNES

24.1 INTRODUCTION

We often perform toxicological research to better understand the mechanism and associated health risk following exposure to hazardous agents. Risk assessment is a systematic scientific characterization of potential adverse health effects following exposure to these hazardous agents. Risk assessment activities are designed to identify, describe, and measure qualities and quantities from these toxicological studies, which are often conducted with homogeneous animal models at doses and exposure duration not encountered in a more heterogeneous human population. Herein lie the challenge of risk assessment. The use of default assumptions because of some level of uncertainty in our extrapolations across species, doses, routes, and interindividual variability, the risk assessment process is often perceived as lacking scientific rigor. This chapter will cover traditional practices as well as new and novel approaches that utilize more of the available scientific data to identify and reduce uncertainty in the process. The advent of powerful computers and sophisticated software programs has allowed the development of quantitative models that better describe the dose-response relationship, refine biologically relevant dose estimates in the risk assessment process, and encourage departure from traditional default approaches (Conolly et al., 1999). Although the focus of this chapter is on current and novel risk assessment methods that are scientifically based, it is critical that the reader be aware of the differences between risk assessment and risk management, which are summarized in Table 24.1.

Results from the risk assessment are used to inform risk management. The risk manager uses the risk information in conjunction with factors such as the social importance of the risk, the social acceptability of the risk, the economic impacts of risk reduction, engineering, and legislative mandates when deciding on and implementing risk management approaches.

The risk assessment may be perceived as the source of a risk management decision, when in fact, social concerns, international issues, trade, public perception, or other non-risk considerations may be taken into consideration. Finally there is one activity known as risk communication that involves making the risk assessment and risk
management information comprehensible to lawyers, politicians, judges, business and labor, environmentalists, and community groups.

### 24.2 RISK ASSESSMENT METHODS

According to the National Research Council of the National Academy of Science, risk assessment consists of four broad but interrelated components: hazard identification, dose-response assessment, exposure assessment, and risk characterization, as depicted in Figure 24.1. The reader should, however, be aware that these risk assessment activities can provide research needs that improve the accuracy of estimating the “risk” or probability of an adverse outcome.

#### 24.2.1 Hazard Identification

In this first component of risk assessment, the question of causality in a qualitative sense in addressed; that is, the degree to which evidence suggests that an agent elicits
a given effect in an exposed population. Among many factors the quality of the studies and the severity of the health effects should be evaluated at this stage. The following are evaluated: (1) validity of the toxicity data, (2) weight-of-evidence summary of the relationship between the substance and toxic effects, and (3) estimates of the generalizability of data to exposed populations. Where there are limited in vivo toxicity data, structural activity relationships (SARs) and short-term assays may be indicative of a chemical hazard. Key molecular structures such as n-nitroso or aromatic amine groups and azo dye structures can be used for prioritizing chemical agents for further testing. SARs are useful in assessing relative toxicity of chemically related compounds, but there are several limitations. For example, toxicity equivalent factors (TEFs) based on induction of Ah receptor by dioxins demonstrated that SARs may not always be predictive. In vitro short-term inexpensive test such as bacterial mutation assays can help identify carcinogens, and there are other short-term tests that can help identify chemicals that potentially can be associated with neurotoxicity, developmental effects, or immunotoxicity. Many of these in vitro studies can provide some insight into mechanism(s) of action, but there may be some false positives and false negatives. Animal studies are usually route-specific and relevant to human exposure, and animal testing usually involves two species, both sexes, 50 animals/dose group, and near-lifetime exposures. Doses are usually 90, 50, and 10 to 25% of the maximum tolerated dose (MTD). In carcinogenicity studies, the aim is to observe significant increases in number of tumors, induction of rare tumors, and earlier induction of observed tumors. However, rodent bioassays may not be predictive of human carcinogenicity because of mechanistic differences. For example, renal tumors in male rats is associated with $\alpha_2\mu$-globulin-chemical binding and accumulation leading to neoplasia; however, $\alpha_2\mu$-globulin is not found in humans, mice, or monkeys. There are differences in susceptibility to aflatoxin-induced tumors between rats and mice that can be explained by genetic differences in expression of cytochrome P450 and GST isoenzymes. Whereas humans may be as sensitive as rats to AFB$_1$-induced liver tumors, mice may not be predictive of AFB$_1$-induced tumors in humans. Epidemiological data from human epidemiological studies are the most convincing of an association between chemical exposure and disease, and therefore can very useful for hazard identification. Exposures are not often well defined and retrospective, and confounding factors such as genetic variations in a population and human lifestyle differences (e.g., smoking) present a further challenge. The three major types of epidemiological studies available are (1) cross-sectional studies, which involve sampling without regard to exposure or disease status, and these studies identify risk factors (exposure) and disease but not useful for establishing cause-effect relationships; (2) cohort studies, which involve sampling on the basis of exposure status, and they target individuals exposed and unexposed to chemical agent and monitored for development of disease, and these are prospective studies; (3) case-control studies, which involve sampling on the basis of disease status. These are retrospective studies, where diseased individuals are matched with disease-free individuals.

**24.2.2 Exposure Assessment**

This process is an integral part of the risk assessment process. However this will be introduced only briefly in this chapter, and the reader is encouraged to consult Chapter 28 in this text as well as numerous other texts that describe the process in
more depth. In brief, exposure assessment attempts to identify potential or completed exposure pathways resulting in contact between the agent and at-risk populations. It also includes demographic analysis of at-risk populations describing properties and characteristics of the population that potentiate or mitigate concern and description of the magnitude, duration, and frequency of exposure. The reader should be aware that exposure may be aggregate (single event added across all media) and/or cumulative (multiple compounds that share a similar mechanism of toxicity). Various techniques such as biomonitoring, model development, and computations can be used to arrive at an estimate of chemical dose taken up by humans, that is, chemical exposure. For example, the lifetime average daily dose (LADD) is a calculation for individuals exposed at levels near the middle of the exposure distribution:

\[
LADD = \frac{(\text{Conc. in media}) \times (\text{Contact rate}) \times (\text{Contact fraction}) \times (\text{Exposure duration})}{(\text{Body weight}) \times (\text{Lifetime})}.
\]

Biological monitoring of blood and air samples represent new ways of reducing uncertainty in these extrapolations. For occupational exposures there are occupational exposure limits (OELs) that are guidelines or recommendations aimed at protecting the worker over their entire working lifetime (40 years) for 8 h/day, 5 days/week work schedule. Most OELs are presented as a time-weighted average concentration for an 8-hour day for a 40-hour work week. There are threshold limit values (TLVs) that refer to airborne concentrations and conditions under which workers may be exposed daily but do not develop adverse health effects. The short-term exposure limit (STEL) are recommended when exposures are of short duration to high concentrations known to cause acute toxicity.

### 24.2.3 Dose Response and Risk Characterization

Dose response is a quantitative risk assessment process, and primarily involves characterizing the relationship between chemical potency and incidence of adverse health effect. Approaches to characterizing dose-response relationships include effect levels such as LD50, LC50, ED50, no observed adverse effect levels (NOAELs), margins of safety, therapeutic index. The dose-response relationship provides an estimation of the relationship between the dose of a chemical agent and incidence of effects in a population. Intuitively, a steep dose-response curve may be indicative of a homogeneous population response, while less steep or almost flat slope may be indicative of greater distribution in response. In extrapolating from relatively high levels of exposure in experimental exposures (usually animals) to significantly lower levels that are characteristic of the ambient environment for humans, it is important to note the shape of the dose-response function below the experimentally observable range and therefore the range of inference. The shape of the slope may be linear or curvilinear and, it should be noted that the focus of risk assessment is generally on these lower regions of the dose-response curve (Figure 24.2).

There is a class of curvilinear dose-response relationships in toxicological and epidemiological studies that may be described as \textit{U-shaped} or \textit{J-shaped curves}. Other terms such as biphasic, and more recently \textit{hormesis}, have been used to refer to paradoxical effects of low-level toxicants. In brief, these dose-response curves reflect an apparent improvement or reversal in the effect of an otherwise toxic agent. These
U-shaped effects can be explained in terms of homeostatic adjustments or overcorrections in the operation of feedback mechanisms. Examples of studies with data fitting a U-shaped curve include the hormetic effect of organic lead on body growth in rats (Cragg and Rees, 1984) and peripheral nerve conduction velocity in children at low doses (Ewert et al., 1986). Similar relationships have been observed with alcohol and nicotine in humans. It has been proposed that because thresholds are inherent in U-shaped dose-response curves, the linear no-threshold extrapolation method is not an appropriate approach for regulating hormetic agents. The current risk assessment paradigm used by US EPA and other federal agencies does not conflict with the concept of hormesis, but it has been proposed that the risk assessor’s analyzes make an active consideration of the data and the application of that data in the low dose portion of the dose-response curve for hormetic agents.

24.3 NONCANCER RISK ASSESSMENT

The noncancer risk assessment process assumes a threshold. For many noncarcinogenic effects, protective mechanisms are believed to exist that must be overcome before an adverse effect is manifested. At the cellular level for some toxicant, a range of exposures exists from zero to some finite value that can be tolerated by the organism with essentially no chance of expression of adverse effects. The aim here in risk assessment is to identify the upper bound of this tolerance range (i.e., the maximum subthreshold level). This approach involves obtaining the no observed adverse effect level. NOAEL is the highest dose level that does not produce a significant elevated increase in an adverse response. Significance refers to biological and statistical criteria and is dependent on dose levels tested, number of animals, background incidence in the unexposed control groups. Sometimes there is insufficient data to arrive at a NOAEL, and a LOAEL (lowest observed adverse effect level) is derived. The NOAEL is the key datum obtained from the study of the dose-response relationship. The NOAEL is used to calculate reference doses (RfD) for chronic oral exposures and reference concentrations (RfC) for chronic inhalation exposures as per EPA. Other agencies, such as the ATSDR and WHO, use the NOAEL to calculate minimum risk levels (MRLs) and acceptable daily intakes (ADI). The US EPA describes the RfD as an estimate, with uncertainty spanning an order of magnitude, of a daily exposure to the human population, including
sensitive subgroups, that is likely to be without appreciable deleterious effects during a lifetime. In deriving reference doses, ADIs, or MRLs, the NOAEL is divided by uncertainty factors (UF) as per EPA (EPA, 1989) and ATSDR (ATSDR, 1993) and by modifying factors (MF) as per EPA:

\[
\text{RfD} = \frac{\text{NOAEL}}{(\text{UF} \times \text{MF})}, \quad \text{US EPA;}
\]

\[
\text{MRL} = \frac{\text{NOAEL}}{\text{UF}}, \quad \text{ATSDR.}
\]

The calculated RfD or RfC is based on the selected critical study and selected critical end point. The risk assessor may obtain numerous studies where the toxicant may have more than one toxic end point, and thus there may be many NOAELs to choose from the literature. In some instances poor data quality may be used to exclude those end points from consideration. Also at issue is the determining what is considered an adverse effect, and this has been summarized with a few examples in Table 24.2. In sum, the MRL or RfD is based on the less serious effects and no serious effects. The following are example effects not used in obtaining a NOAEL: decrease in body weight less than 10%, enzyme induction with no pathologic changes, changes in organ weight with no pathologic changes, increased mortality over controls that is not significant \((p > 0.05)\), and hyperplasia or hypertrophy with or without changes in organ weights.

### 24.3.1 Default Uncertainty and Modifying Factors

Most extrapolations from animal experimental data in the risk assessments require the utilization of uncertainty factors. This is because we are not certain how to extrapolate across species, with species for the most sensitive population, and across duration. To account for variations in the general population and to protect sensitive subpopulations, an uncertainty factor of 10 is used by EPA and ATSDR. The value of 10 is derived from a threefold factor for differences in toxicokinetics and for threefold factor for toxicodynamics. To extrapolate from animals to humans and account for interspecies variability between humans and other mammals, an uncertainty factor of 10 is used by EPA and ATSDR, and as with intraspecies extrapolations, this 10-fold factor is assumed to be associated with toxicodynamics and toxicokinetics. An uncertainty

<table>
<thead>
<tr>
<th>Table 24.2</th>
<th>Comparison of Less Serious Effects and Serious Effects</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Less Serious</strong></td>
<td><strong>Serious</strong></td>
</tr>
<tr>
<td>Reversible cellular changes</td>
<td>Death</td>
</tr>
<tr>
<td>Necrosis, metaplasia, or atrophy</td>
<td>Cancer</td>
</tr>
<tr>
<td>Delayed ossifcation</td>
<td>Clinically significant organ impairment</td>
</tr>
<tr>
<td>Alteration in offspring weight</td>
<td>Visceral or skeletal abnormalities</td>
</tr>
<tr>
<td>Altered T-cell activity</td>
<td>Cleft palate, fused ribs</td>
</tr>
<tr>
<td>Auditory disorders</td>
<td>Necrosis in immunologic components</td>
</tr>
<tr>
<td>50% Reduction in offspring</td>
<td>Visual disorders</td>
</tr>
<tr>
<td></td>
<td>Abnormal sperm</td>
</tr>
</tbody>
</table>
factor of 10 is used when a NOAEL derived from a subchronic study instead of a chronic study is used as the basis for a calculation of a chronic RfD (EPA only). Note that ATSDR does not perform this extrapolation but derive chronic and subchronic MRLs. An uncertainty factor of 10 is used in deriving an RfD or MRL from a LOAEL when a NOAEL is not available. It should be noted that there are no reference doses for dermal exposure, however when there is insufficient dermal absorption data, the EPA uses a default factor of 10% to estimate bioavailability for dermal absorption. A modifying factor ranging from 1 to 10 is included by EPA only to reflect a qualitative professional assessment of additional uncertainties in the critical study and in the entire data base for the chemical not explicitly addressed by preceding uncertainty factors.

Refinements of the RfC have utilized mechanistic data to modify the interspecies uncertainty factor of 10 (Jarabek, 1995). The reader should appreciate that with the inhalation route of exposure, dosimetric adjustments are necessary and can affect the extrapolations of toxicity data of inhaled agents for human health risk assessment. The EPA has included dosimetry modeling in RfC calculations, and the resulting dosimetric adjustment factor (DAF) used in determining the RfC is dependent on physiochemical properties of the inhaled toxicant as well as type of dosimetry model ranging from rudimentary to optimal model structures. In essence, the use of the DAF can reduce the default uncertainty factor for interspecies extrapolation from 10 to 3.16.

The 1996 Food Quality Protection Act (FQPA) now requires that an additional safety factor of 10 be used in the risk assessment of pesticides to ensure the safety of infants and children, unless the EPA can show that an adequate margin of safety is assured with out it (Scheuplein, 2000). The rational behind this additional safety factor is that infants and children have different dietary consumption patterns than adults and infants, and children are more susceptible to toxicants than adults. We do know from pharmacokinetics studies with various human pharmaceuticals that drug elimination is slower in infants up to 6 months of age than in adults, and therefore the potential exists for greater tissue concentrations and vulnerability for neonatal and postnatal effects. Based on these observations, the US EPA supports a default safety factor greater or less than 10, which may be used on the basis of reliable data. However, there are few scientific data from humans or animals that permit comparisons of sensitivities of children and adults, but there are some examples, such as lead, where children are the more sensitive population. It some cases qualitative differences in age-related susceptibility are small beyond 6 months of age, and quantitative differences in toxicity between children and adults can sometimes be less than a factor of 2 or 3.

Much of the research efforts in risk assessment are therefore aimed at reducing the need to use these default uncertainty factors, although the risk assessor is limited by data quality of the chemical of interest. With sufficient data and the advent of sophisticated and validated physiologically based pharmacokinetic models and biologically based dose-response models (Conolly and Butterworth, 1995), these default values can be replaced with science-based factors. In some instances there may be sufficient data to be able to obtain distributions rather than point estimates.

24.3.2 Derivation of Developmental Toxicant RfD

Developmental toxicity includes any detrimental effect produced by exposures during embryonic development, and the effect may be temporary or overt physical malformation. Adverse effects include death, structural abnormalities, altered growth, and
functional deficiencies. Maternal toxicity is also considered. The evidence is assessed and assigned a weight-of-evidence designation as follows: category A, category B, category C, and category D. The scheme takes into account the ratio of minimum maternotoxic dose to minimum teratogenic dose, the incidence of malformations and thus the shape of the dose-response curve or dose relatedness of the each malformation, and types of malformations at low doses. A range of uncertainty factors are also utilized according to designated category as follows: category A = 1–400, category B = 1–300, category C = 1–250, and category D = 1–100. Developmental RfDs are based a short duration of exposure and therefore cannot be applied to lifetime exposure.

**24.3.3 Determination of RfD and RfC of Naphthalene with the NOAEL Approach**

The inhalation RfC for naphthalene was 0.003 mg/m³, and this RfC was derived from a chronic (2-year) NTP inhalation study in mice using exposures of 0, 10, or 30 ppm (NTP, 1992). Groups of mice were exposed for 5 days a week and 6 hours a day. This study identified a LOAEL of 10 ppm. A dose-related incidence of chronic inflammation of the epithelium of the nasal passages and lungs was observed. This LOAEL concentration was normalized by adjusting for the 6-hour-per-day and 5-day-per-week exposure pattern. A LOAEL of 9.3 mg/m³ was obtained was derived by converting 10 ppm first to mg/m³ and then duration-adjusted levels for 6 h/day and 5 days/week for 103 weeks. An UF of 3000 was used, where 10 was for the interspecies (mice to humans) extrapolations, 10 for intraspecies variation in humans, 10 for using a LOAEL instead of a NOAEL, and 3 for database deficiencies.

The oral RfD for naphthalene was 0.02 mg/kg/day, and a study by Battelle (1980) was used to calculate the RfD. Decreased body weight was the most sensitive end point in groups of Fischer 344 rats given 0, 25, 50, 100, 200, or 400 mg/kg for 5 days/week for 13 weeks. These doses were also duration-adjusted to 0, 17.9, 35.7, 71.4, 142.9, and 285.7 mg/kg/day, respectively. The NOAEL for a > 10% decrease in body weight in this study was 71 mg/kg/day. The UF of 3000 was based on 10 for rats to humans extrapolations, 10 for human variation, 10 to extrapolate from subchronic to chronic, and 3 for database deficiencies including lack of chronic oral exposure studies.

**24.3.4 Benchmark Dose Approach**

There are several problems associated with using the NOAEL approach to estimate RfDs and RfCs. The first obvious constraint is that the NOAEL must by definition be one of the experimental doses tested. Once this dose is identified, the rest of the dose-response curve is ignored. In some experimental designs where there is no identifiable NOAEL but LOAEL, the dose-response curve is again ignored, and the NOAEL is derived by application of uncertainty factors as described earlier. This NOAEL approach does not account for the variability in the estimate of the dose response, and furthermore experiments that test fewer animals result in larger NOAELs and thus larger RfDs and RfCs.

An alternative approach known as the benchmark dose (BMD) approach has been developed and implemented by risk assessors as an alternative to the NOAEL approach to estimate RfDs and RfCs. This approach is not constrained by experimental design
as the NOAEL approach, and it incorporates information on the sample size and shape of the dose-response curve. In fact this approach can be used for both threshold and nonthreshold adverse effects as well as continuous and quantal data sets. This requires use of Benchmark Dose Software where the dose-response is modeled and the lower confidence bound for a dose at a specified response level (benchmark response) is calculated. The benchmark response is usually specified as a 1–10% response; that is, it corresponds to a dose associated with a low level of risk such as 1–10%.

Figure 24.3 shows how an effective dose that corresponds to a specific change of effect/response (e.g., 10%) over background and a 95% lower confidence bound on the dose is calculated. The latter is often referred to as the BMDL or LBMD, as opposed to the BMD, which does not have this confidence limited associated with it.

Because the benchmark represents a statistical lower limit, larger experiments will tend, on average, to give larger benchmarks, thus rewarding good experimentation. This is not the case with NOAELs, as there is an inverse relationship between NOAEL and size of experiments. For example, poorer experiments possessing less sensitivity for detecting statistically significant increases in risk inappropriately result in higher NOAELs and RfDs, which may have an unknown unacceptable level of risk. In essence, the NOAEL is very sensitive to sample size, and there can also be high variability between experiments. With the benchmark dose approach, all the doses and slopes of the curve influence the calculations, variability of the data is considered, and the BMD is less variable between experiments. In the BMD approach quantitative toxicological data such as continuous data (organ weights serum levels, etc.) and quantal or incidence data (pathology findings, genetic anomalies, etc.) are fitted to numerous dose-response models described in the literature. The resulting benchmark dose that, for example, corresponds to a tumor risk of 10% generally can be estimated with adequate precision and not particularly dependent on the dose-response model used to fit the data. Note that dose intervals are not required for BMD estimation. This will be greatly appreciated in the cancer risk assessment section of this chapter.

### 24.3.5 Determination of BMD and BMDL for ETU

The BMD method has been quite extensively in assessing quantal data, and very often this has involved analysis of data from developmental and reproductive toxicity

![Figure 24.3 Benchmark dose determination from dose response relationship with the BMDL corresponding to the lower end of a one-sided 95% confidence interval for the BMD.](image)
studies. In this study example (Crump, 1984), rats were exposed to ethylenethiourea (ETU) at 0, 5, 10, 20, 40, and 80 mg/kg doses, and the number affected with fetal anomalies per number of rats were 0/167, 0/132, 1/138, 14/81, 142/178, and 24/24, respectively. The benchmark dose computation can involve utilization of any given dose-response probability model, but in this example the quantal Weibull model was used and the specified effect was set at 0.01 (1%) with confidence level of 0.95. The BMD was determined to be 8.9 mg/kg, and the BMDL was 6.9 mg/kg. This value is close to the NOAEL, which is 5 mg/kg, but it does demonstrate that the NOAEL approximates a lower confidence limit on the BMD corresponding to an excess risk of about 1% for proportions of fetal anomalies. In fact an empirical analysis of some 486 developmental toxicity studies has demonstrated that the NOAEL can result in an excess risk of 5% for proportions of dead or malformed fetuses per litter. The reader should at this stage recognize that the BMD approach can also be used in cancer risk assessment as we are often times working with quantal data that are ideally suited for BMD modeling.

### 24.3.6 Quantifying Risk for Noncarcinogenic Effects: Hazard Quotient

The measure used to describe the potential for noncarcinogenic toxicity to occur is not expressed as the probability. Probabilistic approach is used in cancer RA. For noncancer RA, the potential for noncarcinogenic effects is evaluated by comparing an exposure level (E) over a specified time period with a reference dose (RfD). This ratio is called a hazard quotient:

\[
\text{Hazard quotient} = \frac{E}{RfD}
\]

In general, the greater the value of E/RfD exceeds unity, the greater is the level of concern. Note that this is a ratio and not to be interpreted as a statistical probability.

### 24.3.7 Chemical Mixtures

Human populations are more likely to be exposed simultaneously or sequentially to a mixture of chemicals rather than to one single chemical. Standard default approaches to mixture risk assessment consider doses and responses of the mixture components to be additive. However, it should also be recognized that components in the mixture can also result in synergistic, antagonistic, or no toxicological effect following exposure to a chemical mixture. Therefore mixture toxicity cannot always be predicted even if we know the mechanisms of all toxic components in a defined mixture. Furthermore tissue dosimetry can be complicated by interactions at the route of entry (e.g., GIT, skin surface) and clearance mechanisms in the body. In essence, there are considerable uncertainties involved in trying to extrapolate effects following exposure to chemical mixtures. Several PBPK models have been used to quantitate these effects and also provide some information useful for risk assessment of chemical mixtures (Krishnan et al., 1994; Haddad et al. 2001).

The 1996 FQPA has also mandated that the EPA should also consider implementing cumulative risk assessments for pesticides. Cumulative risk assessments usually involve
integration of the hazard and cumulative exposure analysis, and it primarily involves cumulative nonoccupational exposure by multiple routes or pathways to two or more pesticides or chemicals sharing a common mechanism of toxicity.

Calculation procedures differ for carcinogenic and noncarcinogenic effects, but both sets of procedures assume dose additivity in the absence of information on mixtures:

\[
\text{Cancer risk equation for mixtures: } \text{Risk}_T = \Sigma \text{Risk}_I,
\]

\[
\text{Noncancer hazard index} = \frac{E_1}{\text{RfD}_1} + \frac{E_2}{\text{RfD}_2} + \cdots + \frac{E_i}{\text{RfD}_i}.
\]

This hazard index (HI) approach as well as other indexes (e.g., relative potency factors) are applied for mixture components that induce the same toxic effect by identical mechanism of action. In cases where there are different mechanisms, separate HI values can be calculated for each end point of concern. As the equation above indicates, the HI is easy to calculate, as there is simply scaling of individual component exposure concentrations by a measure of relative potency such as the RfD or RfC, and adding scaled concentrations to get an indicator of risk from exposure to the mixture of concern. However, as noted above, this additivity approach does not take into account tissue dosimetry and pharmacokinetic interactions. Recent published risk assessments have utilized mixture PBPK models to account for multiple pharmacokinetic interactions among mixture constituents. These interaction-based PBPK models can quantify change in tissue dose metrics of chemicals during exposure to mixtures and thus improve the mechanistic basis of mixture risk assessment. Finally the reader should be aware that this HI is different from the a term known as the margin of safety (MOS), which is the ratio of the critical or chronic NOAEL for a specific toxicological end point to an estimate of human exposure. MOS values greater than 100 are generally considered protective if the NOAEL is derived from animal data.

24.4 CANCER RISK ASSESSMENT

For cancer risk assessment an assumption is held that a threshold for an adverse effect does not exist with most individual chemicals. It is assumed that a small number of molecular events can evoke changes in a single cell that can lead to uncontrolled cellular proliferation and eventually to a clinical state of disease. This mechanism is referred to as “nonthreshold” because there is believed to be essentially no level of exposure to such a chemical that does not pose a finite probability, however small, of generating a carcinogenic response. That is, no dose is though to be risk free. Therefore, in evaluations of cancer risks, an effect threshold cannot be estimated. For carcinogenic effects, the US EPA uses a two-part evaluation: (1) the substance is first assigned a weight-of-evidence classification and then (2) a slope factor is calculated.

1. Assigning a weight-of-evidence. The aim here is to determine the likelihood that the agent is a human carcinogen. The evidence is characterized separately for human studies and animal studies as sufficient, limited, inadequate, no data, or evidence of no effect. Based on this characterization and on the extent to which the chemical has been shown to be a carcinogen in animals or humans or both, the chemical is given a provisional weight-of-evidence classification. The US EPA classification system (EPA,
Table 24.3 Weight of Evidence Designation Based on EPA (1986) Guidelines

<table>
<thead>
<tr>
<th>Group</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Human carcinogen</td>
</tr>
<tr>
<td>B1 or B2</td>
<td>Probable human carcinogen</td>
</tr>
<tr>
<td>C</td>
<td>Possible human carcinogen</td>
</tr>
<tr>
<td>D</td>
<td>Not classifiable as to human carcinogenicity</td>
</tr>
<tr>
<td>E</td>
<td>Evidence of noncarcinogenicity for humans</td>
</tr>
</tbody>
</table>

Note: B1 indicates that limited human data are available; B2 indicates sufficient evidence in animals and inadequate or no evidence in humans.

1986) shown in Table 24.3 has been revised in the EPA (1996) proposed guidance and more recent draft guidance (EPA, 1999).

This system was also adapted from the approach taken by the International Agency for Research on Cancer (IARC). This alphanumeric classification system has been replaced with a narrative and the following descriptor categories: known/likely, cannot be determined, or not likely. These EPA (1996) guidelines indicate that not only are tumor findings an important consideration, but also structure-activity relationships, modes of action of carcinogenic agents at cellular or subcellular level and toxicokinetic and metabolic processes. These revised guidelines also indicate that the weighing of evidence should address the conditions under which the agent may be expressed. For example, an agent may “likely” be carcinogenic via inhalation exposure but “not likely” via oral exposure. The narrative will summarize much of this information as well as the mode of action information.

2. Quantifying risk for carcinogenic effects. In the second part of the evaluation, the EPA (1986) guidelines required that quantitative risk be based on the evaluation that the chemical is a known or probable human carcinogen, a toxicity value that defined quantitatively the relationship between dose and response (slope factor) is calculated. Slope factors have been calculated for chemicals in classes A, B1, and B2. Sometimes a value is derived for those in class C on a case-by-case basis. The slope factor is a plausible upper-bound estimate of the probability of a response per unit intake of chemical over a lifetime. Slope factors have been accompanied by the weight-of-evidence classification to indicate the strength of evidence that the chemical is a human carcinogen.

Development of a slope factor entails applying a model to the available data set and using the model to extrapolate from high doses to lower exposure levels expected for human contact. There are a number of low-dose extrapolation models that can be divided into distribution models (e.g., log-probit, Weibull) and mechanistic models (e.g., one-hit, multi-hit, and linearized multistage). EPA 1986 guidelines for carcinogen risk assessment are currently being revised, and it is very likely that the new guidelines will encourage the use of biologically based models for cancer risk assessment. The previous guidelines (EPA, 1986) recommended that the linearized multistage model, which is a mechanistic model, be employed in as the default model in most cases. Most of the other models are less conservative. The proposed biologically based models attempt to incorporate as much mechanistic information as possible to arrive at an estimate of slope factors. In essence, after the data are fit to the selected model, the
upper 95th percent confidence limit of the slope of the resulting dose response curve is calculated. This represents the probability of a response per unit intake over a lifetime, or that there is a 5% chance that the probability of a response could be greater than the estimated value on the basis of experimental data and model used. In some cases, the slope factors based on human dose-response data are based on “best” estimate instead of upper 95th percent confidence limit. The toxicity values for carcinogenic effects can be expressed in several ways.

The slope factor is expressed as $q_1^*$:

$$\text{Slope factor} = \text{Risk per unit dose} = \text{Risk per mg/kg-day}.$$ 

The slope factor can therefore be used to calculate the upper bound estimate on risk ($R$)

$$R = q_1^* [\text{risk} \times (\text{mg/kg/day})^{-1}] \times \text{exposure (mg/kg/day)}.$$ 

Here risk is a unitless probability (e.g., $2 \times 10^{-5}$) of an individual developing cancer and exposure is really chronic daily intake averaged over 70 years: mg/kg/day. This can be determined if we can determine the slope factor and human exposure at the waste site or occupational site. The EPA usually sets a goal of limiting lifetime cancer risks in the range of $10^{-6}$ to $10^{-4}$ for chemical exposures, while the FDA typically aims for risks below $10^{-6}$ for general population exposure. It is therefore quite likely for very high exposures for the accepted EPA range of risk to be exceeded. The EPA range is considered protective of the general and sensitive human population. It should be noted that these orders of magnitude are substantially greater than those used in estimating RfD and RfCs in noncancer risk assessment.

Because relatively low intakes (compared to those experienced by test animals) are most likely from environmental exposure at Superfund hazardous waste sites, it generally can be assumed that the dose-response relationship will be linear on the low-dose portion of the multistage model dose-response curve. The equation above can apply to these linear low-dose situations. This linear equation is valid only at low risk levels (i.e., below the estimated risk of 0.01). For risk above 0.01 the one-hit equation should be used:

$$R = 1 - \exp(-\text{exposure} \times \text{slope factor}).$$

As indicated above, biologically based extrapolation models are the preferred approach for quantifying risk to carcinogens, although it is possible that all the necessary data will not be available for many chemicals. The EPA (1986) guidelines have been modified to include the response data on effects of the agent on carcinogenic processes in addition to data on tumor incidence. Precursor effects and tumor incidence data may be combined to extend the dose response curve below the tumor data; that is, below the range of observation. Thus a biologically based or case-specific dose-response model is developed when there is sufficient data, or a standard default procedure is used when there is insufficient data to adequately curve-fit the data. In brief, the dose-response assessment is considered in two parts or steps, range of observation and range of extrapolation, and the overriding preferred approach is to use the biologically based or case-specific model for both of these ranges. In the first
step of this process, the lower 95% confidence limit on a dose associated with an estimated 10% increase in tumor or nontumor response (LED₁₀) is identified. When human real world exposures are outside the range of the observed or experimental data, this serves as the point of departure or marks the beginning for the extrapolating to these low environmental exposure levels. Note that these procedures are very similar to the benchmark procedure for quantitating risk to noncancerous chemicals. In the second step, the biologically based or case-specific model is preferred for use in extrapolations to lower dose levels provided that there are sufficient data. If the latter is not the case, then default approaches consistent with agent chemical mode of action are implemented with the assumption of linearity or nonlinearity of the dose-response relationship. The linear default approach is a departure from the 1986 guidelines, which used the linearized multistage (LMS) procedure, but is based on mode of action or alternatively if there is insufficient data to support a nonlinear mode of action. In brief, it involves drawing a straight line from the point of departure (LED₁₀) to the origin (i.e., zero). When there is no evidence of linearity or there is a nonlinear mode of action, the default approach is the margin of exposure (MOE) analysis. The MOE approach computes the ratio between the LED₁₀ and the environmental exposure, and the analysis begins from the point of departure that is adjusted for toxicokinetic differences between species to give a human equivalent dose.

Finally it should be noted that prior to the FQPA in 1996, the Delaney clause prohibited the establishment of tolerances or maximum allowable levels for food additives if it has been shown to induce cancer in human or animal. This is an important change in regulations because pesticide residues were considered as food additives. Because of the FQPA, pesticide residues are no longer regarded as food additives, and there is no prohibition against setting tolerances for carcinogens.

24.5 PBPK MODELING

Physiologically based pharmacokinetic (PBPK) modeling has been used in risk assessment to make more scientifically based extrapolations, and at the same time to help explore and reduce inherent uncertainties. Historically pharmacokinetics has relied on empirical models, and in many instances this process offers little insight into mechanisms of absorption, distribution, and clearance of hazardous agents and does not facilitate translation from animal experiments to human exposures. For example, dose scaling using by body weight or size may often time overestimate or underestimate toxicant levels at the target tissue. PBPK models can help predict tissue concentrations in different species under various conditions based on independent anatomical, physiological, and biochemical parameters. In these analyzes physiological parameters such as organ volumes, tissue-blood partition coefficients, and blood flow to specific tissue compartments described by the model, are calculated or obtained from the literature and integrated into the model. Monte Carlo analysis, a form of uncertainty analysis, can now be performed, and this allows for the propagation of uncertainty through a model that results in estimation of the variance of model output. This can be achieved by randomly sampling model parameters from defined distributions; some parameters such as cardiac output, metabolic, and log P parameters, may have a lognormal distribution, while other parameters may be normal or uniform. In essence, the Monte Carlo analysis when coupled with PBPK characterizes the distribution of potential risk
in a population by using a range of potential values for each input parameter (not single values) as well as an estimate of how these values are distributed (Clewell and Andersen, 1996). By these approaches, uncertainty is identifiable and quantifiable, and can reduce inappropriate levels of concern in reporting the risk of chemical exposure. These mathematical modeling approaches also help identify areas of potential scientific research that could improve the human health assessment.

In recent years there have been significant efforts at harmonizing noncancer and cancer risk assessments (Barton et al., 1998; Clewell et al., 2002), and in this respect PKPD modeling can be a very useful tool in the risk assessment process. For example, recall that noncancer risk assessment addresses variability in a population by dividing the NOAEL by 10, whereas the cancer risk assessment does not address this quantitatively. PBPK modeling coupled with Monte Carlo analysis is one approach as described in the previous paragraph that will help address this level of uncertainty in the risk assessment. In conclusion, it should be noted that PBPK modeling has been utilized with very few toxicants. It is hoped that risk assessment policy will encourage the use of this tool as well as other appropriate models to integrate mechanistic information and the pharmacokinetics (dosimetry), and pharmacodynamics (dose response) of toxicants. Improved quantitative risk assessments will ultimately provide scientifically sound information that will influence the risk management decision process.

SUGGESTED READING


