TOXICOLOGICAL SCIENCES 61, 32–39 (2001)
Copyright © 2001 by the Society of Toxicology

FORUM

Applications of Mechanistic Data in Risk Assessment: The Past, Present, and Future

Lynne T. Haber,[1] Andrew Maier, Qiyu Zhao, Joan S. Dollarhide, Russell E. Savage, and Michael L. Dourson

Toxicology Excellence for Risk Assessment (TERA), 1757 Chase Avenue, Cincinnati, Ohio 45223

Received October 23, 2000; accepted January 12, 2001

Mechanistic data, when available, have long been considered in risk assessment, such as in the development of the nitrate RfD based on effects in a sensitive group (infants). Recent advances in biology and risk assessment methods have led to a tremendous increase in the use of mechanistic data in risk assessment. Toxicokinetic data can improve extrapolation from animals to humans and characterization of human variability. This is done by the development of improved tissue dosimetry, by the use of uncertainty factors based on chemical-specific data, and in the development of physiologically based pharmacokinetic (PBPK) models. The development of the boron RfD illustrates the use of chemical-specific data in the improved choice of uncertainty factors. The draft cancer guidelines of the U.S. Environmental Protection Agency emphasize the use of mode of action data. The first choice under the guidelines is to use a chemical-specific, biologically based dose-response (BBDR) model. In the absence of a BBDR model, mode of action data are used to determine whether low-dose extrapolation is done using a linear or nonlinear (margin of exposure) approach. Considerations involved in evaluating a hypothesized mode of action are illustrated using 1,3-dichloropropene, and use of a BBDR model is illustrated using formaldehyde. Recent developments in molecular biology, including transgenic animals, microarrays, and the characterization of genetic polymorphisms, have significant potential for improving risk assessments, although further methods development is needed. Overall, use of mechanistic data has significant potential for reducing the uncertainty in assessments, while at the same time highlighting the areas of uncertainty.

Key Words: **risk assessment; mode of action; mechanism; polymorphism; nitrate; boron; telone; 1,3-dichloropropene; formaldehyde.**

Although noncancer and cancer risk assessments conducted in the past 20 years have sometimes used mechanistic data, the use of such data is increasing as toxicologists and risk assessors develop a better understanding of how chemicals exert their toxic effects. For the purposes of this paper, "mechanistic data" refers to all data other than that obtained from classical toxicity testing (e.g., toxicokinetics data, data on precursor effects, and studies on the interactions of the chemical of interest with cellular macromolecules). Use of mechanistic data can lead to a better choice of critical effect, identification of a sensitive population, and/or to higher confidence in the hazard identification and dose response assessment. Properly used, all of these applications of mechanistic data improve the quality of the resulting assessment and reduce the associated uncertainty.

Risk assessors now routinely consider mechanistic and kinetic data in the development of noncancer assessments, particularly in the extrapolation of data from animals to humans and in the selection of uncertainty factors (UFs) in calculating reference doses (RfDs), reference concentrations (RfCs), or acceptable daily intakes (ADIs). Such considerations appeared in the late 1980's in the development of RfCs by the U.S. Environmental Protection Agency (U.S. EPA) with the advent of dosimetric adjustments for converting chemical concentrations inhaled by experimental animals to human equivalent concentrations (Jarabek, 1994, 1995; U.S. EPA, 1994). The associated modifications to UFs were later captured more formally by Renwick (1993), who proposed a data-derived UF approach. In this latter approach, both the UF for extrapolation from animals to humans and the UF for interindividual variability are divided into subfactors to allow for separate evaluations of differences in toxicokinetics and toxicodynamics. Modifications of this approach have been used by Health Canada (Meek *et al.*, 1994) and the International Programme on Chemical Safety (IPCS, 1994). As default values, the IPCS suggests that the UF for interspecies extrapolation should be subdivided unequally into 4-fold (toxicokinetics) and 2.5-fold (toxicodynamics), and that the UF for intraspecies extrapolation be split evenly (3.16-fold for both toxicokinetics and toxicodynamics). These splits were based on the distribution of inter- and intraspecies differences in toxicokinetics and toxicodynamics among a small number of pharmaceuticals (Ren-

[1] To whom correspondence should be addressed. Fax: (513) 542-7487. E-mail: haber@tera.org.

wick, 1993) and were supported by a later evaluation of a larger group of pharmaceuticals (Renwick and Lazarus, 1998). When chemical-specific data are available, these categorical defaults may be replaced by compound-specific adjustment factors (CSAFs) (Meek *et al.*, in press). The choice of UFs can also take into account whether the effects of the chemical progress with continued exposure (e.g., subchronic vs. chronic), whether there is evidence for differential susceptibility of adults and children, and whether the chemical is an essential element. The use of CSAFs and consideration of these issues is illustrated in the case study on boron.

Recent guidelines and methods currently under development expand the use of mechanistic data for cancer risk assessments. For example, one of the major ways in which the 1996 draft of the U.S. EPA cancer guidelines differs from the 1986 guidelines is in the consideration of mode of action (MOA) data. Mode of action is not explicitly defined in the guidelines, but it can be considered to be a general description of the manner in which a chemical acts to produce its toxic effect. Knowing the mode of action means that the obligatory precursor step in the carcinogenic process has been identified. This precursor step may involve, for example, a chemical interacting directly with DNA and forming DNA adducts, or the chemical causing cytotoxicity, leading to compensatory cell proliferation. Under the 1996 cancer guidelines, the preferred approach for low-dose extrapolation is using a biologically based dose-response (BBDR) model. In the absence of such a model, mode of action understanding is used in determining whether extrapolation to low doses is done using a linear extrapolation, a nonlinear one, or both procedures. Use of such considerations in risk assessment is illustrated in the case studies for formaldehyde and 1,3-dichloropropene. A recent interagency initiative, led by the U.S. EPA, is expanding on the mode of action framework proposed by the U.S. EPA 1996 cancer guidelines. This framework includes the development of dosimetric approaches for the oral, inhalation, and dermal routes, taking into account the chemical's dosimetry, site of toxic action including the portal of entry, and its mode of action (Jarabek *et al.*, manuscript in preparation).

In this review, we highlight how the application of mechanistic data is having an important impact on the derivation of risk estimates, while also pointing out opportunities for the further evolution of the basic science and the tools for application of these data. To accomplish this, we have selected case studies from recent risk assessments as a backdrop for discussing current applications of mechanistic data. The role that mechanistic data plays in improved tissue dosimetry, application of early biological effects data, and replacement of default UFs with data-derived ones are discussed within the context of the case studies. We further discuss current data gaps, with the goal of communicating to basic biology researchers our view of the data needs of risk assessment scientists.

Case Studies

Nitrate: identification of sensitive population. One of the earliest applications of mechanistic data was in the use of information about the sensitive population to move away from a default UF for the nitrate RfD. Ingestion of drinking water contaminated with nitrate from fertilizer runoff can lead to methemoglobinemia in infants and possibly in adults. Nitrate is converted to nitrite in the gut and absorbed into the bloodstream, where it chemically oxidizes hemoglobin to methemoglobin (MetHb). Unlike hemoglobin, MetHb is unable to reversibly bind oxygen. Infants are much more sensitive than children or adults to this effect, for four reasons:

- Infants have a less acidic gastric environment, and therefore their guts have high levels of nitrate-reducing bacteria, which survive only under less acidic conditions.
- Infants under 3 months of age also have a lower enzymatic capacity of MetHb reductase, the enzyme that reduces MetHb back to hemoglobin.
- Infants have a form of hemoglobin (fetal hemoglobin) that is particularly susceptible to oxidation by nitrite.
- Infants' water intake per body weight is higher than that of adults, thus contributing to higher dose per kilogram of body weight.

The RfD for nitrate is based on the NOAEL for clinical signs of methemoglobinemia in infants (U.S. EPA, 2000). Because the NOAEL was identified in a sensitive subgroup of the population, a 1-fold UF for within-human variability in toxicokinetics and toxicodynamics was used to estimate the RfD in place of the default value of 10. In this case, a full understanding of the mechanism of nitrate toxicity in the sensitive subgroup led to a high confidence RfD that was equivalent to the NOAEL of the critical effect in the sensitive subgroup.

Boron: use of toxicokinetic data in the choice of UFs. Boron compounds are widely distributed in the environment. The major sources of human exposure are food and drinking water; fruits, vegetables, and nuts are naturally rich in boron. Boron is also used in laundry detergent and in household pesticides. In animals, boron exposure results in reproductive and developmental effects; decreased fetal weight is the most sensitive effect. The critical NOAEL of 9.6 mg/kg/day was reported by Price *et al.* (1996), based on decreased fetal weight. In light of the extensive information available on boron toxicokinetics in humans and animals, a number of authors have used this end point, together with data-derived UFs, to develop exposure level guidance (e.g., IPCS, 1994; the World Health Organization's Working Group on Chemical Substances in Drinking Water (WGCSDW); Murray and Andersen, in press). All of these assessments used the same NOAEL (or the related benchmark dose) as the basis for the assessment, although the UFs ranged from 25 to 60. This discussion focuses on the WGCSDW assessment (published as Dourson *et al.*, 1998) and the Murray and Andersen (in press)

update, as being the most recent assessments and best representing the state of the science.

Over a wide range of doses, studies in both experimental animals and humans have found that ingested boron is almost completely absorbed (summarized in Dourson *et al.*, 1998). Boron distributes evenly throughout the body fluid by passive diffusion. In rats and humans, boron concentrations were comparable in almost all tissues examined, except that there were somewhat higher concentrations in bone (Alexander *et al.*, 1951; Forbes *et al.*, 1954; Ku *et al.*, 1991; Shuler *et al.* 1990; Ward, 1993). Metabolism of inorganic borate by biological systems is not thermodynamically feasible due to the excessive energy required to break the boron–oxygen bond (Emsley, 1989). Based on the quantitative similarity in absorption, distribution, and metabolism, no UF was needed to account for species differences in these aspects of toxicokinetics. However, based on the steady-state blood boron concentrations and the corresponding oral doses, boron clearance rates were estimated to be 40 ml/kg/h in humans, and 163 ml/kg/h in rats. This indicates that rats have an approximately 3- to 4-fold higher boron clearance rate than humans. Based on these considerations, the WGCSDW (Dourson *et al.*, 1998) proposed a factor of 4 for interspecies differences in elimination. More recent work found that boron clearance in pregnant rats is 3.1 times as fast as in pregnant women when renal clearance is normalized to body weight (milliliters/minute/kilogram), supporting a factor of 3.1 for interspecies differences in toxicokinetics (Murray and Andersen, in press). Because no data were available to support a UF other than 2.5 for interspecies variation in toxicodynamics, the WGCSDW used a default value of 2.5 for the toxicodynamic variation, which together with the boron-specific value of 4 for toxicokinetic differences resulted in an interspecies UF of 10. Murray and Andersen (in press) proposed a reduced interspecies factor for toxicodynamics of 1.25–2.5, based on interspecies similarities in toxicodynamics for a number of end points, and based on the nutritional importance, if not essentiality, of boron. A total UF of 3.9–7.8 for interspecies extrapolation resulted using this alternative approach.

The intraspecies UF for toxicokinetic variability was also based solely on variation in excretion. Because no information on boron clearance in pregnant women was available, WGCSDW estimated the variability in boron clearance rate in pregnant women based on the available information on variability in the glomerular filtration rate (GFR) during human pregnancy. Considering two standard deviations from the mean GFR as a reasonable measure of variability in light of the biological lower limit of GFR consistent with good health, a factor of 1.8 resulted. Murray and Andersen (in press) reported on a recent study of boron clearance in pregnant women, which found that a two standard deviation variance from the mean in pregnant women translates to a 2.0-fold difference in clearance, a value close to the 1.8-fold value used by the WGCSDW. Because there are no data on human variability in toxicodynamics, both of the assessments used the default UF of 3.2 ($10^{0.5}$) for human toxicodynamics. Thus, both groups used subfactors in the range of 1.8–2 for the variation in toxicokinetics and a default value of 3.2 for the variation in toxicodynamics, resulting in a UF of 6 for intraspecies variation. This case study illustrates the need for criteria for determining data-derived UFs, in light of the variety of approaches previously used to determine the safe dose of boron. The development of such criteria has been the focus of recent IPCS activity (Meek *et al.*, in press). However, it is noteworthy that the composite UFs calculated by the various different groups are within a factor of 2 to 3. Additional research is needed to consider the issue of how the risk assessment should be modified if the chemical of interest is an essential element.

1,3-Dichloropropene: use of mode of action data on glutathione (GSH) depletion in cancer risk assessment. 1,3-dichloropropene is the active ingredient in commercially applied preplanting fumigants used for the control of nematodes. Recent inhalation cancer risk assessments for 1,3-dichloropropene (U.S. EPA, 2000) and Telone II, an epoxidized soybean oil stabilized 1,3-dichloropropene formulation (Dow AgroSciences, Indianapolis; manuscript in preparation; meeting notes from a peer review of the Dow assessment are available at http://tera.org/peer/) highlight how mechanistic toxicity data can have a direct impact on quantitative risk estimates. In both of these assessments, increased incidence of bronchioalveolar adenoma in male mice was used as the basis of the cancer assessment. The assessments differed, however, in the interpretation of the evidence for a MOA supporting a nonlinear extrapolation to low doses. In the Dow assessment, both margin of exposure (MOE) and linear low-dose modeling approaches were presented, whereas the U.S. EPA assessment concluded that a linear low-dose modeling approach was more appropriate.

Support for the application of a nonlinear assessment hinged on the determination of whether GSH depletion, resulting in a decreased ability to detoxify the administered dose, caused the tumorigenic response seen in mice. The Dow assessment argued that the absence of detection of potentially DNA-reactive epoxide metabolites in the presence of detoxification enzymes and reduced GSH, coupled with generally negative results in *in vivo* genotoxicity assays, supports the conclusion that 1,3-dichloropropene lacks genotoxic potential under normal physiological conditions. The Dow assessment further suggested that the reported genotoxic activity in some *in vivo* and *in vitro* studies (using the current soybean oil stabilized product formulation) is associated with doses that saturate detoxification pathways. As support for this conclusion, the Dow assessment cited the concordance between doses that deplete GSH and those that induce tumors in animal studies. As further support for a nongenotoxic MOA, the Dow assessment noted that Telone II induced tumors only late in life and of a type having high spontaneous incidence.

The U.S. EPA assessment (U.S. EPA, 2000) also examined the data in support of GSH depletion as a key event in the MOA, but concluded that a linear low-dose approach was most appropriate. This conclusion was based on the mixed results in mutagenicity assays, coupled with the absence of adequate support for the proposed MOA, as described below. The assessment (U.S. EPA, 2000) noted that "1,3-dichloropropene may be nongenotoxic at low-dose exposures that do not interfere significantly with normal function of GSH, but bioassay data showing the protective effect of GSH against tumor formation are lacking." The language in the U.S. EPA assessment suggests that the proposed GSH-depletion mechanism would support a nonlinear risk assessment approach, given sufficient evidence for its causal role in tumor induction. This being the case, what data would be needed to build a sufficient case to support GSH depletion as a key event in the tumorigenic MOA?

The following evidence for a role of GSH depletion in the lung tumor induction is provided to highlight the use of mechanistic data in support of identification of an MOA. Lomax *et al.* (1989) reported a concentration-dependent increase in the incidence of bronchioalveolar adenomas in male B6C3F1 mice exposed to 1,3-dichloropropene (Telone II) for 2 years; the tumor incidences were 9/50, 6/50, 13/50, and 22/50 in mice exposed to 0, 5, 20, and 60 ppm Telone II, respectively. To test whether this might be due to GSH depletion, W. T. Stott *et al.* in an unpublished study (described in greater detail in U.S. EPA, 2000) measured lung GSH levels in male mice exposed by inhalation for up to 26 days over the tumorigenic dose range. GSH was decreased in a concentration-dependent manner after 3, 12, and 26 days of dosing. The mean lung GSH levels were 85.3, 79.9, 58.3, and 43.1% of control levels following exposures to 10, 30, 60, or 150 ppm, respectively, after 26 days of exposure. In parallel samples collected during the 26-day inhalation study, no changes in cell proliferation or apoptosis rates were observed, indicating that these latter two events are unlikely to be related to 1,3-dichloropropene-induced tumorigenesis.

Reviewing these data under the framework of the 1996 cancer guidelines (U.S. EPA, 1996), as further explicated in the January 1999 SAB draft revision of the cancer guidelines (U.S. EPA, 1999), provides an opportunity to highlight the types of mechanistic studies that would more fully support the proposed MOA. The U.S. EPA framework focuses on identification of key events in the development of tumors. A key event is defined as an empirically observed precursor consistent with a mode of action. According to these guidelines, considerations in weighing the evidence for a proposed MOA include magnitude, consistency, and specificity of association; the dose-response relationship and temporal relationship between the suggested key event and tumor induction; and biological plausibility and coherence with known biology of the disease.

The data are limited for evaluation of consistency between lung GSH depletion and induction of the lung adenomas, primarily because each of these end points has been examined *in vivo* in only a single study in a single species; a bioassay has been conducted in rats without observing lung adenomas. The data also appear to be limited regarding the specificity of the response, as GSH depletion is a common cellular response to many toxicants, not necessarily all of which are tumorigenic. In an unpublished study by W. T. Stott *et al.* (1997), GSH depletion was measured only in tissues that developed tumors. Evidence showing an absence of GSH depletion in tissues that do not have a tumorigenic response would lend support for the specificity of the key event. Together, the issues of specificity and consistency address whether the key event of GSH depletion is a necessary event for 1,3-dichloropropene–induced tumor development. It is noteworthy that GSH depletion may not be a sufficient event, but that does not rule out the hypothesized MOA. While the available data are not inconsistent with an effect of GSH depletion, the proposed MOA identifies only a single key event. A stronger case could be made for the role of GSH homeostasis if additional intermediate events that link GSH depletion to tumorigenesis were also demonstrated experimentally. Concentrations of 1,3-dichloropropene that effectively depleted GSH had no effect on two plausible intermediate effects, cell proliferation or apoptosis (W. T. Stott *et al.*, 1997, unpublished data). Thus, the link between cellular thiol status and cell growth needs to be further clarified. For example, studies in mice given cotreatments with GSH synthesis inhibitors should yield an enhanced tumorigenic response, whereas coadministration of antioxidants should decrease 1,3-dichloropropene tumorigenicity. Additional mechanistic toxicology studies can provide relevant approaches for more tightly linking GSH depletion to tumorigenesis.

The strength of the dose-response and temporal relationships between key events and the outcome of the tumorigenic response are also to be considered in the weight of evidence evaluation. Both GSH depletion and the induction of lung adenomas were dose dependent. In addition, GSH depletion was significantly affected at doses lower than those associated with adenoma formation, and the degree of GSH depletion increased with increasing dose. Therefore, the dose-response data are consistent with GSH depletion being an early event in lung tumor formation. These data would be consistent with GSH depletion being necessary, but not sufficient, for lung tumor formation, although this consistency alone is not sufficient to show causality. Lung GSH depletion was measured after 26 days of exposure, whereas lung adenoma incidence increased after 2 years of dosing; continued GSH depletion throughout a significant period of tumorigenesis has not been shown. Nonetheless, because GSH depletion occurred within 3 days of dosing and was maintained for up to 26 days with daily dosing (W. T. Stott *et al.*, 1997, unpublished data), GSH depletion may have continued with longer exposure durations, suggesting a temporal relationship consistent with a role of GSH depletion. Thus, there are data supporting the hypothesized relationship between GSH depletion and tumor development, but additional

data are needed to meet criteria of the U.S. EPA for sufficiency to support a nonlinear low-dose extrapolation.

Formaldehyde: use of a biologically based dose-response model for cancer risk assessment. As noted above, the 1996 cancer guidelines (U.S. EPA, 1996) recommend that a BBDR model be used for low-dose extrapolation if sufficient data are available. BBDR models incorporate mechanistic data into the risk assessment process by providing a mathematical description of the link between an exposure metric and the key mechanistic events that drive downstream toxic responses. The use of dose-response data for key events in the mechanism of toxicity provides the opportunity to more accurately predict the risk of low-dose exposures, as dose-response information is developed directly for precursor events, which may occur at or near environmental exposure levels. As for other applications of mechanistic data, a positive demonstration must be made that the key events modeled are necessary steps in the development of the adverse effect of interest.

BBDR modeling has been described in the context of both cancer and noncancer risk assessments. Application of BBDR modeling was incorporated into a recent assessment of formaldehyde cancer risk (CIIT, 1999). A BBDR model has also been described for noncancer end points such as chloroform hepatotoxicity (Conolly and Butterworth, 1995). BBDR modeling efforts for developmental toxicity of 5-fluorouracil have also been described recently (Lau *et al.*, 2000).

The recent inhalation cancer risk assessment for formaldehyde provides a detailed example of the application of BBDR modeling (CIIT, 1999). Although the epidemiology data provide mixed results, formaldehyde is a nasal tumorigen in rats following exposure by the inhalation route (Kerns *et al.*, 1983; Monticello *et al.*, 1996; Morgan *et al.*, 1986; reviewed in the CIIT 1999 assessment). The CIIT assessment (CIIT, 1999) used a two-stage clonal growth model of carcinogenesis that incorporated data on the rate of cell division and on the number of cells at risk in different regions of the respiratory tract of rats and humans. Regional uptake of inhaled formaldehyde was simulated in each species, and the flux values were determined for the upper respiratory tracts of rats and humans using three-dimensional, anatomically accurate computational fluid dynamics (CFD) models of rat, primate, and human nasal airflow and inhaled gas uptake (Kepler *et al.*, 1998; Kimbell *et al.*, 1997; Subramaniam *et al.*, 1998; reviewed in CIIT, 1999).

Two types of key events in tumorigenesis were considered: the formation of DNA–protein crosslinks (DPX) or a related genotoxic event, and increased mutation fixation through regenerative cell proliferation secondary to cytotoxicity. A human BBDR model was developed. Predicted dose-response relationships for DPX formation from rat and monkey data were used to develop clonal growth model parameters related to the direct mutagenic component of the model. The labeling index dose-response in rats (a measure of regenerative cell proliferation, normalized by unit length) was used to represent the cell proliferation component of the clonal growth model. Other model parameter estimates were derived from human lung cancer incidence data or were taken from the published literature. The resulting two-stage clonal growth model was subsequently used to estimate cancer risk in humans under various exposure scenarios.

The application of the BBDR modeling approach results in low-dose risk estimates that are much lower than those obtained from approaches that rely only on tumor incidence data and apply default linear low-dose extrapolations. For example, in the CIIT formaldehyde assessment, the cancer risk estimates for humans following exposure to 0.1 ppm 6 h/day, 5 days/week was compared for different risk assessment methods. The human risk was reported as 2.7E-8 based on BBDR modeling and 4.2E-4 based on benchmark dose modeling of tumor incidence data and linear low-dose extrapolation.

The appropriateness of a BBDR model, and thus the accuracy of the low-dose risk estimates, depends on the degree of evidence that the low-dose end points that drive the shape of the dose-response curve are really key events in the MOA. Gaps in the data on the relationship between the precursor events and the outcome of interest result in uncertainty in the assessment. The CIIT assessment (CIIT, 1999) indicates that the application of the model reduces overall uncertainty, as sources of uncertainty are identified and the degree of uncertainty can be incorporated into the modeling. A number of key areas of uncertainty in the development of the clonal growth model were noted. For the use of DPX, data gaps that were discussed include defining the relationship between DPX and the induction of mutations, determining the kinetics of DPX formation and removal, and evaluating the appropriateness of assumptions for interspecies extrapolation. An uncertainty noted in the use of the unit length labeling index (ULLI) to model effects on cell proliferation was the assumption that the toxicodynamics for this end point was similar in the human and rat nasal cells. An additional uncertainty was the use of estimated cell cycle times for the respiratory epithelium, rather than measured species-specific values.

The case of formaldehyde highlights the potential utility of BBDR modeling for cancer risk assessment. The application of BBDR models provides the opportunity to reduce uncertainty in risk estimates across the entire exposure range of interest, primarily by using information on tissue dose of the toxic moiety, and by incorporating dose-response information for precursor events. As with the more qualitative margin of exposure approach for moving away from the default for low-dose extrapolation, BBDR modeling requires that an adequate positive demonstration be made that the proposed key events in the MOA are necessary steps in the cancer process induced by the exposure of interest.

Future Directions

Dose-response determinations for early biological effects. The case studies highlighted above for 1,3-dichloropropene and formaldehyde provide examples where preclinical effect

measures played an important part in drawing conclusions about the likely shape of the dose-response curve. The use of data on early effects is attractive for this purpose, because much of the uncertainty surrounding risk assessments results from the need to predict risk at environmentally relevant doses, which are often below dose levels that induce clearly adverse effects in animal studies. Biomarkers of early effects offer the possibility of reducing many of the uncertainties associated with low-dose extrapolation by extending the range of observation to include validated response end points in place of less sensitive defaults (morbidity, mortality, tumors, etc.). Considerations for the use of biomarkers of early effects in the dose-response analysis have been outlined elsewhere (Hattis and Silver, 1993; NRC, 1987; Schulte, 1989; U.S. EPA, 1994), and critical evaluations of the utility of effect biomarkers and further research needs have been discussed in several more recent reviews (Aitio and Kallio, 1999; Groopman and Kensler, 1999; Mutti, 1999; Ward and Henderson, 1996).

As demonstrated in the case studies, early-effect biomarkers can have an important impact on the outcome of the risk assessment. However, use of biomarkers in risk assessment depends on the adequate characterization of the link between the biomarker and other events in the exposure–disease continuum, using criteria such as those originally outlined by Hill (1965) and expanded on by Schulte *et al.* (1989) and U.S. EPA (1994). Some key elements of the Hill criteria that must be demonstrated for determining whether a key precursor event has been identified include biological plausibility, and evaluation of how the dose-response and time-response relationships of the precursor event relate to those of the adverse effect of interest (U.S. EPA, 1994, 1999). Laboratory scientists can enhance the relevance of their studies to risk assessment by including dose-response evaluation (ideally tied to the doses used for toxicity or carcinogenicity studies) in their mechanistic studies. Another way to enhance the utility of their studies includes ensuring that the models chosen are relevant to the key toxic effects (e.g., the target tissue is evaluated).

Several newer approaches can be used to link precursor events to downstream toxicity end points and to test the strength of that linkage. New tools in molecular biology offer opportunities for testing whether an early event is an obligate precursor to more traditional outcomes. Transgenic animals constitute such a new tool and include animals with a targeted disruption of a gene of interest (knockout) as well as animals with other targeted gene changes. Bull (2000) recently described the potential for using data from a knockout mouse strain to provide critical evidence regarding whether mouse liver tumors induced by trichloroacetic acid (TCA) can be attributed to peroxisome proliferation. As suggested in that article, the absence of TCA-induced liver tumors in mice with a targeted disruption of the gene encoding the peroxisome proliferation receptor would be a powerful demonstration that activation of this receptor is a key event in TCA-induced tumor formation.

The capacity to detect molecular and cellular events in response to toxicants is rapidly increasing, as exemplified by the measurement of toxicant-induced gene expression changes using DNA microarrays (e.g., Nuwaysir *et al.*, 1999). Microarray data differ from single gene expression experiments by simultaneously measuring gene expression changes representing gene clusters that reflect a range of toxic severity. For example, a DNA array provides information on the induction of gene clusters involved in adaptive responses as well as cell proliferation or cytotoxicity. Conducting a dose-response experiment using DNA arrays yields a series of dose-response curves for each gene cluster. By defining the quantitative relationships among the dose-response relationships of these different gene clusters, it may be possible to elucidate the dose-response relationship between adaptive and adverse effects.

A traditional approach for demonstrating the relationship among early effect biomarkers and more adverse end points is to test their correlation using regression approaches. Categorical regression is a meta-analytical technique that allows for modeling of the severity of response versus the dose (Dourson *et al.*, 1997; Guth *et al.*, 1997; Teuschler *et al.*, 1999). This approach might offer an alternative to test the link between dose-response curves for a series of effects of increasing severity, to aid in biomarker validation. Based on this scheme, if the precursor event is truly linked to more severe events at higher doses, then inclusion of the dose-response data for the precursor end point with the dose-response data for increasingly severe end points in the same continuum would result in a good fit in the categorical regression model. When these criteria are met, the categorical regression model provides an opportunity to expand the dose-response curve to include the lower doses represented by the precursor events.

Dose-response information for precursor end points may also be useful for extrapolating to environmentally relevant exposures and for improving animal-to-human extrapolation. The methods for using precursor end points for these purposes have not yet been fully developed, but are active areas of research. BBDR modeling is likely to provide useful insights, as such modeling defines the relationship between tissue dose and defined biological steps. This was done for formaldehyde, as described above.

It is noteworthy that quantitatively linking the dose-response data for precursor and clearly adverse events can be problematic. For example, based on a 2,3,7,8-tetrachlorodibenzo-*p*-dioxin case study, Andersen and Barton (1998) suggested that precursor effects, such as CYP1A1 and CYP1A2 induction, are frequently not proportionally related to the toxic response of interest, such as tumorigenic responses. This is due in part to the contribution of multiple mechanisms in the development of the downstream pathogenesis. Under these conditions, direct application of the precursor dose-response curve in the risk assessment can often be used to yield a conservative risk estimate (Andersen and Barton, 1998).

It may also be possible to use precursor data to improve low-dose extrapolation without having the detailed mechanistic information needed to include precursor information in a BBDR model. Precursor end points may be measurable in exposed human populations, but risk assessors need to determine the relationship between those measured precursors and the adverse effect of interest. If the dose-response relationship between the precursor and tissue dose in humans is determined, and the quantitative relationship between a precursor effect and the adverse effect of interest can be determined in animals, also on the basis of tissue dose, a prediction of the human dose response for adverse effects can be made. Categorical regression is a promising approach for developing such a quantitative connection between the precursor and adverse effect, because it allows the precursor and adverse effects to be modeled in either an interdependent or independent manner. A limitation to this approach, however, is that it assumes that the relationship between precursor and adverse effects is the same in the experimental animal species and in humans.

Using polymorphism data to characterize intraspecies variability. As described in the introduction and in the boron case study, there has been considerable recent work on using chemical-specific data in place of the default UFs for extrapolating from animals to humans and for accounting for human variability. One of the important sources of human variability in susceptibility to toxicant responses is genetic polymorphisms (Taningher *et al.,* 1999; Wormhoudt *et al.*, 1999). Unfortunately, methods have not been developed for the routine incorporation of data on polymorphisms into risk assessments. Some initial attempts have been conducted to address this issue (e.g., Clewell and Andersen, 1996; Renwick and Lazarus, 1998). However, further investigation is needed to determine the relevance of this type of variation in the estimation of intraspecies UFs as applied in risk assessments.

One promising approach for quantifying the effect of genetic polymorphisms is through the combined use of physiologically based pharmacokinetic (PBPK) models and Monte Carlo analysis. Use of this approach allows one to determine how variability in all relevant parameters determines the variability in the ultimate tissue dose. Because it is unlikely that an individual will be at the edge of the distribution for multiple unrelated physiological parameters, incorporating all sources of variability may result in a smaller total variability than that expected from considering only one parameter (such as enzyme polymorphisms). Similarly, if enzyme activity is not rate limiting at environmental exposure levels, enzyme polymorphisms may not have a significant effect on tissue dose.

Summary

Mechanistic data, when available, have long been considered in risk assessment, but the use of such data is increasing with the recent explosion of mechanistic studies, molecular biology techniques, and mathematical approaches for addressing mechanistic issues. Use of such data has significant potential for reducing the uncertainty in assessments, while at the same time making the areas of uncertainty more explicit.

REFERENCES

Aitio, A., and Kallio, A. (1999). Exposure and effect monitoring: A critical appraisal of their practical application. *Toxicol. Lett.* **108,** 137–147.

Alexander, G. V., Nusbaum, R. E., and MacDonald, N. S. (1951). The boron and lithium content of human bones. *J. Biol. Chem.* **192,** 489–496.

Andersen, M. E., and Barton, H. A. (1998). The use of biochemical and molecular parameters to estimate dose-response relationships at low levels of exposure. *Environ. Health Perspect.* **106**(Suppl. 1), 349–355.

Bull, R. J. (2000). Mode of action of liver tumor induction by trichloroethylene and its metabolites, trichloroacetate and dichloroacetate. *Environ. Health Perspect.* **108**(Suppl. 2), 241–259.

CIIT (1999). Formaldehyde: Hazard characterization and dose-response assessment for carcinogenicity by the route of inhalation. Chemical Industry Institute of Toxicology, Research Triangle Park. Available online at http://www.ciit.org/newsrs/formaldehydesummary. Accessed September, 2000.

Clewell, H. J., and Andersen, M. E. (1996). Use of physiologically-based pharmacokinetic modeling to investigate individual versus population risk. *Toxicology* **111,** 315–329.

Conolly, R. B., and Butterworth, B. E. (1995). Biologically based dose response model for hepatic toxicity: A mechanistically based replacement for traditional estimates of noncancer risk. *Toxicol. Lett.* **82/83,** 901–906.

Dourson, M. L., Maier, A., Meek, B., Renwick, A., Ohanian, E., and Poirier, K. (1998). Boron tolerable intake: Re-evaluation of toxicokinetics for data-derived uncertainty factors. *Biol. Trace Elem. Res.* **66,** 453–463.

Dourson, M. L., Teuschler, L. K., Durkin, P. R., and. Stiteler, W. M. (1997). Categorical regression of toxicity data: A case study using aldicarb. *Regul. Toxicol. Pharmacol.* **25,** 121–129.

Emsley, J. (1989). *The Elements.* Clarendon Press, Oxford.

Forbes, R. M., Cooper, A. R., and Mitchell, H. H. (1954). On the occurrence of beryllium, boron, cobalt and mercury in human tissues. *J. Biol. Chem.* **209,** 857–865.

Groopman, J. D., and Kensler, T. W. (1999). The light at the end of the tunnel for chemical-specific biomarkers: Daylight or headlight? *Carcinogenesis* **20,** 1–11.

Guth, D. J., Carroll, R. J., Simpson, D. G., and Zhou, H. (1997). Categorical regression analysis of acute exposure to tetrachloroethylene. *Risk Anal.* **17,** 321–332.

Hattis, D., and Silver, K. (1993). Use of biomarkers in risk assessment. In *Molecular Epidemiology: Principles and Practice* (P. A. Schulte and F. P. Perera, Eds.), pp. 251–273. Academic Press, San Diego.

Hill, A. B. (1965). The environment and disease: Association or causation? Proc. R. Soc. Med. **58,** 295–300.

IPCS (1994). Environmental health criteria 170: Assessing human health risks of chemicals: Derivation of guidance values for health-based exposure limits. International Programme on Chemical Safety, World Health Organization, Geneva.

Jarabek, A. M. (1994). Inhalation RfC methodology: Dosimetric adjustments and dose-response estimation of noncancer toxicity in the upper respiratory tract. *Inhal. Toxicol.* **6**(Suppl), 301–325.

Jarabek, A. M. (1995). The application of dosimetry models to identify key processes and parameters for default dose-response assessment approaches. *Toxicol. Lett.* **79,** 171–184.

Kepler, G. M., Richardson, R. B., Morgan, K. T., and Kimbell, J. S. (1998). Computer simulation of inspiratory nasal airflow and inhaled gas uptake in a rhesus monkey. *Toxicol. Appl. Pharmacol.* **150,** 1–11.

Kerns, W. D., Pavkov, K. L., Donofrio, D. J., Gralla, E. J., and Swenberg, J. A. (1983). Carcinogenicity of formaldehyde in rats and mice after long-term inhalation exposure. *Cancer Res.* **43,** 4382–4392.

Kimbell, J. S., Godo, M. N., Gross, E. A., Joyner, D. R., Richardson, R. B., and Morgan, K. T. (1997). Computer simulation of inspiratory airflow in all regions of the F344 rat nasal passages. *Toxicol. Appl. Pharmacol.* **145,** 388–398.

Ku, W. W., Chapin, R. E., Moseman, R. F., Brink, R. E., Pierce, K. D., and Adams, K. Y. (1991). Tissue disposition of boron in male Fischer rats. *Toxicol. Appl. Pharmacol.* **111,** 145–151.

Lau, C., Andersen, M. E., Crawford-Brown, D. J., Kavlock, R. J., Kimmel, C. A., Knudsen, T. B., Muneoka, K., Rogers, J. M., Setzer, R. W., Smith, G., and Tyl, R. (2000). Evaluation of biologically based dose-response modeling for developmental toxicity: A workshop report. *Regul. Toxicol. Pharmacol.* **31,** 190–199.

Lomax, L. G., Stott, W. T., Johnson, K. A., Calhoun, L. L., Yano, B. L., and Quast, J. F. (1989). The chronic toxicity and oncogenicity of inhaled technical grade 1,3-dichloropropene in rats and mice. *Fundam. Appl. Toxicol.* **12,** 418–431.

Meek, M. E., Newhook, R., Liteplo, R. G., and Armstrong, V. C. (1994). Approach to assessment of risk to human health for priority substances under the Canadian Environmental Protection Act. *Environ. Carcinogen. Ecotoxicol. Rev.* **C12,** 105–134.

Meek, B., Renwick, A., Ohanian, E., Dourson, M., Naumann, B, and Vu, V. (2001). Guidelines for application of compound specific adjustment factors (CSAF) in dose/concentration response assessment. *Comments Toxicol.* (in press).

Monticello, T. M., Swenberg, J. A., Gross, E. A., Leininger, J. R., Kimbell, J. S., Seilkop, S., Starr, T. B., Gibson, J. E., and Morgan, K. T. (1996). Correlation of regional and nonlinear formaldehyde-induced nasal cancer with proliferating populations of cells. *Cancer Res.* **56,** 1012–1022.

Morgan, K. T., Jiang, X. Z., Starr, T. B., and Kerns, W. D. (1986). More precise localization of nasal tumors associated with chronic exposure of F-344 rats to formaldehyde gas. *Toxicol. Appl. Pharmacol.* **82,** 264–271.

Murray, F. J., and Andersen, M. E. (2001). Data-derived uncertainty factors: Boric acid (BA) as a case study. *Hum. Ecol. Risk Assess.* (in press).

Mutti, A. (1999). Biological monitoring in occupational and environmental toxicology. *Toxicol. Lett.* **108,** 77–89.

NRC (1987) National Research Council. Biological Markers in Environmental Health Research. *Environ. Health Perspect.* **74,** 3–9.

Nuwaysir, E. F., Bittner, M., Trent, J., Barrett, J. C., and Afshari, C.A. (1999). Microarrays and toxicology: The advent of toxicogenomics. *Mol. Carcinog.* **24,** 153–159.

Price, C. J., Strong, P. L., Marr, M. C., Meyers, C. B., and Murray, F. J. (1996). Developmental toxicity NOAEL and postnatal recovery in rats fed boric acid during gestation. *Fundam. Appl. Toxicol.* **32,** 179–193.

Renwick, A. G. (1993). Data derived safety factors for the evaluation of food additives and environmental contaminants. *Food Addit. Contam.* **10,** 275–305.

Renwick, A. G., and Lazarus, N. R. (1998). Human variability and noncancer risk assessment–an analysis of the default uncertainty factor. *Regul. Toxicol. Pharmacol.* **27,** 3–20.

Schulte, P. A. (1989). A conceptual framework for the validation and use of biologic markers. *Environ. Res.* **48,** 129–144.

Shuler, T. R., Pootrakul, P., Yamsukon, P., and Nielson, F. H. (1990). Effect of thalassaemia/immunoglobin E disease on macro, trace and ultratrace element concentrations in human tissues. *J. Trace Elem. Exp. Med.* **3,** 31–43.

Subramaniam, R. P., Richardson, R. B., Morgan, K. T., Kimbell, J. S., and Guilmette, R. A. (1998). Computational fluid dynamics simulations of inspiratory airflow in the human nose and nasopharynx. *Inhal. Toxicol.* **10,** 92–120.

Taningher, M., Malacarne, D., Izzotti, A., Ugolini, D., and Parodi, S. (1999). Drug metabolism polymorphisms as modulators of cancer susceptibility. *Mutat. Res.* **436,** 227–261.

Teuschler, L. K., Dourson, M. L., Stiteler, W. M., McClure, P., and Tully, H. (1999). Health risk above the reference dose for multiple chemicals. *Regul. Toxicol. Pharmacol.* **30,** S19–S26.

U.S. EPA (2000). Integrated Risk Information System (IRIS). U. S. Environmental Protection Agency. Online, www.epa.gov/iris. Accessed September, 2000.

U.S. EPA (1999). Guidelines for carcinogen risk assessment. NCEA-F-0644. Review draft. U.S. Environmental Protection Agency, Washington, D.C.

U.S. EPA (1996). Proposed Guidelines for Carcinogen Risk Assessment. EPA/600/P-92/003C. Office of Research and Development, U.S. Environmental Protection Agency, Washington, D.C.

U.S. EPA (1994). Methods for Derivation of Inhalation Reference Concentrations and Application of Inhalation Dosimetry. EPA/600/8-90-066F. Office of Health and Environmental Assessment, U.S. Environmental Protection Agency, Washington, D.C.

Ward, N. I. (1993). Boron levels in human tissues and fluids. In *Trace Elements in Man and Animals-TEMA 8. Proceedings of the Eighth International Symposium on Trace Elements in Man and Animals* (M.A.D. Meissner and C.-F. Mills, Eds.), pp. 724–728. Verlag Media Touristik, Gersford, Germany.

Ward, J. B., and Henderson, R. E. (1996). Identification of needs in biomarker research. *Environ. Health Perspect.* **104**(Suppl. 5), 895–900.

WHO (1998). *Environmental Health Criteria: 204 Boron.* World Health Organization, International Programme on Chemical Safety, Geneva.

Wormhoudt, L. W., Commandeur, J. N. M., and Vermeulen, N. P. E. (1999). Genetic polymorphisms of human N-acetyltransferase, cytochrome P450, glutathione-*S*-transferase, and epoxide hydrolase enzymes: Relevance to xenobiotic metabolism and toxicity. *Crit. Rev. Toxicol.* **29,** 59–124.