Human and Ecological Risk Assessment: Vol. 8, No. 6, pp. 1421-1444 (2002)

Improving Risk Assessment: Research Opportunities in Dose Response Modeling to Improve Risk Assessment[1]

Lauren Zeise,[2] Dale Hattis,[3] Mel Andersen,[4] A. John Bailer,[5] Steve Bayard,[6] Chao Chen,[7] Harvey Clewell,[8] Rory Conolly,[9] Kenny Crump,[8] David Dunson,[10] Adam Finkel,[11] Lynne Haber,[12] Annie M. Jarabek,[13] Ralph Kodell,[14] Daniel Krewski,[15] Duncan Thomas,[16] Todd Thorslund,[6] and James T. Wassell[17]

ABSTRACT

Substantial improvements in dose response modeling for risk assessment may result from recent and continuing advances in biological research, biochemical techniques, biostatistical/mathematical methods and computational power. This

Report of the "Dose Response Models" Breakout Group, of the Workshop "Future Research for Improving Risk Assessment Methods: Of Mice, Men, and Models," August 16-18, 2000, Snowmass, Colorado, sponsored by the National Institute for Occupational Safety and Health, National Institute of Environmental Health Sciences, US Environmental Protection Agency, Chemical Manufacturers Association (currently American Chemistry Council), and United Auto Workers. Invited participants of the workgroup are the authors of this paper. Additional participants and contributers to the discussion were: Ronald Brown, James Deddens, Brian Miller, Robert Park, Raghupathy Ramanathan, Jan Sassaman, Edward Slavin, Randall Smith, Robert Stenner, Jana Sorvari, and David Utterback.

[2]California Environmental Protection Agency, Office of Environmental Health Hazard Assessment, Oakland, California; breakout group chair. [3]Research Professor, Marsh Institute, Clark University, 950 Main Street, Worcester, Massachusetts 01610, (p) (508) 751-4603, (f) (508) 751-4600, dhattis@aol.com; breakout group rapporteur. [4]Colorado State University, Fort Collins, Colorado. [5]Miami University, Oxford, Ohio. [6]US Occupational Safety and Health Administration; Washington, DC, submitted research proposal but was unable to attend workshop. [7]US Environmental Protection Agency, Office of Research and Development, National Center for Environmental Assessment, Washington, DC. [8]K.S. Crump Group/ICF Consulting, Ruston, Louisiana. [9]CIIT Centers for Health Research (previously Chemical Industry Institute of Toxicology), Research Triangle Park, North Carolina. [10]National Institute of Environmental Health Sciences, Research Triangle Park, North Carolina. [11]US Occupational Safety and Health Administration; Washington, DC. [12]Toxicology Excellence for Risk Assessment, Cincinnati, Ohio. [13]US Environmental Protection Agency, Office of Research and Development, National Center for Environmental Assessment, Research Triangle Park, North Carolina. [14]National Center for Toxicological Research, Jefferson, Arkansas. [15]University of Ottawa, Ottawa, Ontario. [16]University of Southern California, Los Angeles, CA. [17]National Institute for Occupational Safety and Health, Morgantown, WV.

1080-7039/02/$.50
© 2002 by ASP

report provides a ranked set of recommendations for proposed research to advance the state of the art in dose response modeling. The report is the result of a meeting of invited workgroup participants charged with identifying five areas of research in dose response modeling that could be incorporated in a national agenda to improve risk assessment methods. Leading topics of emphasis are interindividual variability, injury risk assessment modeling, and procedures to incorporate distributional methods and mechanistic considerations into now-standard methods of deriving a reference dose (RfD), reference concentration (RfC), minimum risk level (MRL) or similar dose-response parameter estimates.

Key Words: interindividual variability, injury risk, Bayesian analysis, latent variables, Markov Chain Monte Carlo analysis.

INTRODUCTION

The recent advances in biological research, biochemical techniques, molecular epidemiology, biostatistical/mathematical methods, and computational power provide opportunities for considerable improvement in the assessment of dose response relationships for risk assessment. This paper describes the discussion and results of a one-day meeting of thirteen invited participants (listed as authors) of a dose response workgroup and several additional attending experts in dose response modeling. The workgroup meeting was part of a three-day workshop exploring research possibilities to improve risk assessments done to support occupational and environmental health policies and standards. The charge to the participants of the workgroup was to identify five specific research areas that would significantly improve dose response modeling for occupational and environmental risk assessment. Each invited participant was asked to prepare a brief description of research programs that could be the subject of a request for proposals. Table 1 shows titles and authors of each initial written proposal. This paper draws from the proposal write-ups, often verbatim, to describe the areas of research the workgroup found most promising.

The dose-response workgroup meeting proceeded in three stages. First proposals were sorted into broad topic areas, and then each proposal was presented and discussed. Additional topic areas and ideas for proposals were then solicited from the group. Next there was an extended period during which the individual proposals were further refined, extended, and consolidated, into ten groups. Finally, each participant was asked to rank the proposal groups giving five points to the top choice in terms of desirability for funding, four for the second choice, three for the third, two for the fourth, one for the fifth and zero for the remaining five proposal groups. In all, thirteen people participated in this ranking exercise.

PROPOSED LINES OF DOSE RESPONSE RESEARCH

The ten areas the group chose to explore, in ranked order are given in Table 2. Some proposals benefited by the multiplicity of ideas attached within the various groups—and the process of grouping is the principal reason why there is not a one-to-one mapping of the initial proposals listed in Table 1 to the final ranked items in Table 2. The groupings may well have been a disadvantage to some of the

Table 1. Initial research ideas proposed to improve dose response modeling for risk assessment.

Participant	Research Proposal
Mel Andersen	Mechanistic Models for the Impact of Lifestyle Factors on Metabolism, Pharmacokinetics, and Toxicity of Workplace Chemicals in Diverse Human Populations
John Bailer	Developing Strategies for Quantitative Risk Estimation for Hazards of Occupational Injury
Chao Chen	Advances in Molecular Research and Risk Assessment
Harvey Clewell	Accounting for Pharmacokinetic Uncertainty and Variability in Risk Assessment
Rory Conolly	Risk Assessment for What?
Kenny Crump	Investigation of Prevalence of Hormesis
David Dunson	Risk Assessment Based on Multiple Reproductive and Developmental Endpoints
Lynne Haber	1. Use of Data-Derived Uncertainty Factors in Human Risk Assessment 2. Characterizing Uncertainty in RfDs and RfCs
Annie Jarabek	1. Characterizing Variability to Decrease Uncertainty 2. Extending the Range of Observation
Ralph Kodell	Development of Mechanistic Biomathematical Models that Predict Less-than-Background Risk of Cancer at Low Doses of Toxicants
Duncan Thomas	Incorporating Genetics into Risk Assessment
Todd Thorslund	Proposed Research to Strengthen the Rationale for Extrapolating
Steve Bayard	Animal Cancer Bioassay Results to Humans Based on Relative Risk
James Wassell	Occupational Injury Risk Assessment

proposals that might have benefited from being joined with others, but were presented in relative isolation from other proposals. In any event, the differences between the second through the seventh ranked proposal groups are relatively small, and the eighth proposal also received significant support. This paper discusses the first eight ranked groups, rather than the five top rated proposals originally requested by the organizers.

CHARACTERIZING INTERINDIVIDUAL AND INTERSPECIES VARIABILITY IN SUSCEPTIBILITY

An important consideration in regulators' decisions to mitigate risks is that people differ in their responses to environmental and occupational exposures. During the last 15 years, innovations in both observational study design and risk analysis methodology have led to improved descriptions of interindividual differences in risk (see, *e.g.*, Bogen and Spear 1987; Bois *et al.* 1996 and 1999; Hattis and Burmaster 1994; Cullen and Frey 1999) and in the mechanistic bases of pharmacokinetic and pharmacodynamic determinants of exposure-dose-response across species used in laboratory testing (USEPA 1994; Jarabek 1995a,b; USEPA 1996; Schlosser and Bogdanffy 1996). Yet, mathematical tools to reach policy decisions that adequately address characterization of exposure-dose-response in a mechanistic fash-

Table 2. Summary ranking of ideas for the dose response modeling workgroup.

1. Characterizing Interindividual and Interspecies Variability in Susceptibility
2. Models for Injury Risk Assessment – Including Development and Characterization of Risk Estimates and Exposure Metrics
3. Adaptations and Modifications to Existing Standard Procedures (*e.g.*, to Derive RfDs)
4. Acceptability Criteria for Mechanistic Hypotheses and Data
5. Modeling Analyses for Multiple Endpoint Data (Especially for Different Endpoints on Different Scales)
6. New Mechanistic Models of Carcinogenesis
7. Combining Data of Different Types in Risk Analyses
8. Exploring Evidence and Models for Complex Dose Response Relationships in the Context of Homeostasis
9. Ideals for Risk Analysis/ Relationship with Societal Decision Processes
10. Interspecies Extrapolation Based on Relative Risk

ion or that address heterogeneity have not been adopted by regulatory agencies for routine use. The standard practice is to base standards on the population risk, sometimes in combination with ad hoc procedures thought to be conservative (*e.g.*, use of most sensitive species) to estimate an upper bound on population risk. It is hoped that decisions based on an upper bound so calculated will provide protection for sensitive individuals. To make such estimates, simple dose-time-response or stochastic models of the carcinogenesis process are fit to experimental or epidemiological data. Noncancer methods rely on designation of a sentinel adverse effect in a putatively sensitive species, and then apply interspecies and intrahuman "uncertainty factors" that have been traditionally based on empirical, not mechanistic, motivation (Jarabek 1995a).

Break-out group participants identified several lines of research that would lead to improved methods to assess interindividual variability in dose response. These are outlined below.

A. Methods to Incorporate Genetic Determinants of Heterogeneity to Model Variability in Cancer Risk

It is generally accepted that cancer is a genetic disease, involving somatic mutations or other changes to DNA that can be induced by environmental exposures to carcinogens. The mathematical models used in risk assessment such as the Armitage-

Doll multistage model and the Moolgavkar-Knudson two-mutation clonal expansion model are based on fundamental concepts about mutation and heterogeneity. However, neither model accounts for an individual's specific genetic make-up, or population heterogeneity in unmeasured genetic or environmentally-induced factors. Some germline mutations can also produce a hereditary predisposition to cancer or an unusual sensitivity to environmental carcinogens. For example, Ataxia-Telangiectasia patients (homozygous for the *ATM* mutation) are exquisitely sensitive to ionizing radiation, and *ATM* heterozygotes are at elevated risk. Other examples include predisposing genes such *BCRA1* and *BCRA2* for familial breast and ovarian cancer, *APC* for familial polyposis coli, and the mismatch repair genes *MLH1* and *MSH2* for hereditary nonpolyposis colorectal cancer. Polymorphisms in genes involved in the metabolic activation of pre-carcinogens to their active form or their deactivation, also may confer a larger population attributable risk (e.g., *NAT1* and *NAT2* activation of aromatic amines). A number of these polymorphisms affect appreciable fractions of the population. Finally genomic instability is an important mechanism in certain kinds of cancer, such as hereditary nonpolyposis colon cancer, in which mutations in one gene produces extensive loss of DNA replication fidelity, leading to high rates of somatic mutation at other loci involved in the carcinogenic process. Absent specific information about an individual's genetic makeup, substantial heterogeneity between individuals in their baseline risk and sensitivity to environmental carcinogens is expected.

Duncan Thomas proposed development of mathematical methods to model heterogeneity that would take advantage of the powerful molecular tools currently available for genetic testing (*e.g.*, including whole genomic scans and gene-expression arrays). Priorities noted for possible development included:

1. Methods for incorporating variation in identifiable genetic factors and for estimating residual heterogeneity

2. Stochastic models of carcinogenesis incorporating genomic instability. The key idea is to allow for the possibility that an early event in carcinogenesis could lead to a somatic mutation in a mismatch repair or other regulatory gene, inducing a complex cascade of subsequent events. The data from microdissected tumors on clonal variation in molecular markers within a single tumor could potentially be exploited for this purpose.

3. Methods for incorporating information on metabolic genes into physiologically-based pharmacokinetic models to describe complex pathways (elaborated further below).

Expanding on these ideas, the group recognized the value of methods exploiting data for specific metabolic gene polymorphisms, induction of activating and detoxifying enzymes, DNA repair, and major gene changes on defined germ-line mutations in genes along known molecular pathological pathways in cancer. The quantitative work of Hattis and Barlow (1996) based on variability in phenotype rather than genotype observations (activities of metabolic activation, inactivation, and DNA repair) could be extended through greater use of in vivo observations of relevant enzyme activities, and procedures to separate underlying variability from

measurement errors (see, *e.g.*, Hattis and Silver 1994). Methods to measure variability in both susceptibility and relevant exposures by examining the pattern of age-specific cancer incidence could be further developed. For example, Finkel (1987, 1995) and others (Manton and Stallard 1979; Manton *et al.* 1986) using techniques of heterogeneity dynamics have extracted estimates of variability from age specific incidence data. Heterogeneity dynamics provide methods for describing changing characteristics of a heterogeneous population as its members age.

B. Characterizing Interindividual Variability in Effective Doses and Risk within the Framework of Physiologically Based Pharmacokinetic Models

Occupational risks from chemicals vary among individuals in a workplace due to varying exposure patterns and differences among workers' rates for enzymatic activation and detoxification of chemicals, health status, and other factors. Enzyme activity is determined by various intrinsic and extrinsic factors. These include genetic makeup, lifestyle, pharmaceutical usage, alcohol intake, health status, and age. For example, drugs serve as potential enzyme inducers. Dietary factors, health status (*e.g.*, ketotic states associated with diabetes), and alcohol alter the activity of an important oxidative enzyme, P450-2E1, involved in metabolism of some workplace chemicals. The impacts of complicated, time-dependent interactions of multiple factors on metabolism, and the complex exposure patterns in the workplace, are often overlooked in the setting of occupational standards.

Melvin Andersen proposed research to investigate conditions under which interactions are likely to enhance toxicity, and to identify members of the population most at risk. The research would explore the quantification of increased risk, and the identification of activities and lifestyle factors that lead to higher workplace risks. How such characterizations could influence PEL or TLV standard setting activities for occupational exposures would also be examined.

Past assessments of interactions have mostly involved observation of altered pharmacokinetic behavior of compounds in healthy human volunteers under controlled exposure conditions. Conclusions were typically drawn from observed alterations in the kinetic properties, for example, in the presence of chemical mixtures, with intake of small to moderate amounts of alcohol, or with varying levels of exercise. It is proposed that methods now be developed utilizing physiologically based toxicokinetic (PBTK) and toxicodynamic (TD) models. PBTK models describe the disposition, metabolism and transport of chemicals and metabolites in various tissues of the body, and certain TD models characterize alteration of enzyme levels, receptor levels, and effects due to exogenous compounds or dietary factors (see, *e.g.*, Chien *et al.* 1997). The PBTK/TDs modeling framework provides the means for integrating critical data of different types into the assessment. This includes data from mechanistic studies of pharmacokinetics and enzyme induction in animals, and limited studies in human volunteers or with human tissue. The overall goal of the modeling research would be to define the temporal and lifestyle factors that lead to variability in responses to chemical exposure in diverse worker populations, and to quantify that variability.

Variability in risk due to interindividual differences in pharmacokinetics has been addressed in recent risk assessments, including that by the Occupational Safety

and Health Administration for methylene chloride, the US Environmental Protection Agency's draft assessment for trichloroethylene, and California's draft assessment for tetrachloroethylene (Cal/EPA OEHHA, 2000, based on Bois *et al.* 1996). Recent uses of PBTK models have also evaluated the adequacy of current standard setting approaches when variability is taken into account (Thomas *et al.* 1999).

In a related proposal, Harvey Clewell proposed research to investigate the hierarchical Bayesian approach to characterizing uncertainty and variability in pharmacokinetic models for cancer risk assessment, as developed by Frederic Bois and others (Bois *et al.* 1996; Gelman *et al.* 1996). The power of PBTK models is obtained at the expense of using a large number of parameters, some of which may vary significantly among individuals (*e.g.*, the pharmacokinetic constants) and few of which are known with precision. The impact of parameter uncertainty in PBTK models has often been evaluated using a Monte Carlo approach (Bois *et al.* 1990; Allen *et al.* 1996; Clewell and Andersen 1996; Bailer and Dankovic 1997; Dankovic and Bailer 1994; Hattis 1990; Portier *et al.* 1989), wherein specified probability distributions are randomly sampled for each model parameter and the PBTK model is run. The process is repeated numerous times to define a probability distribution for the desired PBPK dose metric. In typical applications of this approach, variability in individual toxicodynamic response susceptibility per unit of internal dose is not addressed, and if it is, it is not decoupled from uncertainty (*e.g.*, due to measurement error). Further, although model parameters are correlated, only limited correlation is typically assumed, or it is ignored altogether.

Hierarchical statistical models within a Bayesian framework have been applied to disentangle model uncertainty from variability, using the computational technique of Markov chain Monte Carlo simulation (Bois 1999; Bernillon and Bois 2000). Markov chain Monte Carlo (MCMC) procedures are widely useful for fitting of hierarchical models that incorporate individual-specific parameters. For example, MCMC algorithms have been used for estimating the population distribution of physiological parameters based on data for the pharmacokinetics of the chemical in different individuals (Bois *et al.* 1996; Jonsson and Johanson 2000). The approach adopts a hierarchical population model to enable uncertainty and variability in an individual's response to be distinguished from the variability of individual responses within the population. The same pharmacokinetic model structure applies to all individuals, but model parameters vary among individuals. The Bayesian framework provides a formal structure for combining prior knowledge on parameters from the scientific literature with data from pharmacokinetic experiments, to generate posterior distributions for any given parameter value. Thus widely different types of data can be integrated, for example, from studies of distribution and elimination in human volunteers, in vitro and in vivo metabolic studies in experimental animals, and physiological measurements from various sources. The overall approach provides the statistical foundation to support PBPK model calibration that is lacking in most PBPK applications (Bernillon and Bois 2000; Kohn 1995 and 1997).

Some members of the group raised concerns over cases when the use of this approach results in parameter estimates that differ substantially from values expected based on a priori knowledge. This may occur when the "prior" distributions assumed are relatively broad, and parameter values are significantly influenced by distantly related parameters with extensive observations used in the Bayesian "up-

dating". The programs and analyses are complicated, and there is the concern regarding how possible subtle misspecification of model structure and possible underestimation of the uncertainty in available observational data might influence posterior estimates of parameter values. The proposed study is, using both actual and simulated data, to distinguish the conditions under which reestimation of model parameters with the approach produces relatively accurate results from those under which it may be misleading. This might be done, for example, by applying the approach to a well-characterized system, deleting information, and predicting the deleted data from the remaining information (a technique referred to as "cross validation").

Other proposed work in this area would focus on greatly needed research on model uncertainty and lack of identifiability in complex models. Both may be dealt with effectively using a Bayesian approach. For example, if a biologically realistic model is under identified, prior information can be brought in to enable Bayesian identifiability. In addition, if there are definable uncertainties in the model, one can utilize Bayesian model averaging techniques.

C. Human Variability in Baseline Values for Parameters as Predictors of Non-Cancer Susceptibility

One important determinant of the population distribution of susceptibility to non-cancer toxic insults is the baseline distribution in the human population of functions and functional reserve capacities for physiological process such as kidney, lung, or liver functions. We define functional reserve capacity as the amount of change in a physiological parameter needed to produce abnormal function or an adverse outcome. As an example from Hattis *et al.* (1999) the distribution of low density lipoprotein (LDL) is considered. LDL is thought to be an important physiological parameter likely to be on the causal pathway to cardiovascular disease. The variability of LDL in the population loosely reflects differing susceptibility to cardiovascular disease. The distribution of functional capacity among those not receiving intervention might be described as the distribution of differences between a standard cutoff for clinical intervention and measured values in the population. Baseline observation studies, such as surveying via NHANES LDL levels, have the advantage that they do not require deliberate administration of toxicants or drugs to humans.

From LDL and other parameters one can predict the risk of cardiovascular disease. Similarly, from study of other continuous variables related to serious health outcomes it may be possible to develop relationships for use in predicting non-cancer risks. Examples include, birth weight and infant mortality; sperm quality parameters and male fertility performance; forced expiratory volume in one second as a predictor for general cardiovascular mortality; iodine deficiency and thyroid pathology. Also by examining the distribution of indicators of functional capacity one can gauge the extent to which the certain risk assessment practices are protective (*e.g.*, assignment of certain uncertainty factors). Dale Hattis proposed the development of data to explore the use of baseline observations in quantitative non-cancer risk assessment procedures. For additional discussion of the potential for this type of study, see Hattis (1998) and Hattis *et al.* (1999).

D. Variability in Mechanistic Determinants of Chemical Disposition (*e.g.* Related to Age, Species, Sex, Disease State, for Oral, Inhalation, and Dermal Exposures)

As discussed above, dosimetry based on PBTK models has become a useful tool to adjust for differences in delivered and internal dose across species and within human populations. Examples of applications in risk assessments for gases include: formaldehyde (CIIT 1999), tetrachloroethylene (Bois *et al.* 1996), vinyl acetate (Bogdanffy *et al.* 1999), EGBE (IRIS 1999), and vinyl chloride (IRIS 2000). Dosimetry modeling was also a key aspect of the National Ambient Air Quality Standard for particulate matter in 1996a (USEPA 1996). Reduced mechanistic model structures and empirical models of mass transport have formed the basic default procedures for the calculation of human equivalent concentrations in the U.S. Environmental Protection Agency reference concentration methods (USEPA 1994; Jarabek 1995b).

Annie Jarabek proposed research to fill substantive gaps in the anatomical, physiological and mechanistic data needed to explicitly describe the major factors influencing chemical disposition defined to encompass the processes of deposition (*e.g.* primary deposition of an inhaled toxicant or particles on airway surfaces), uptake, distribution, metabolism and elimination as well as subsequent toxicant-target interactions, *i.e.*, mode of action (Jarabek 2000). Such data would allow a comprehensive characterization of the exposure-dose-duration-response continuum across species so that variability can be addressed by describing differences in the mechanistic factors that determine disposition and pathogenesis. She proposed the acquisition of such data for different ages and genders in experimental animals and humans and for different disease states (*e.g.*, COPD) in humans for inhalation, oral and dermal routes.

Parameters of interest include the anatomical parameters of airway lengths, portal of entry (respiratory tract, GI, dermis) tissue thickness, cell types and locations, and other physiological parameters such as ventilation rates, GI transport rates, dermal transport rates, metabolism, and the fraction of specific types of cells in various phases of the cell cycle. These data would support improved dosimetry modeling to increase the accuracy of descriptions of dose differences among species and within the human population. For example, the respiratory tract of children differs dramatically from that of the adult in anatomical structure and ventilatory pattern (*e.g.*, the oral-nasal switching point between nose versus mouth-breathing with exertion) and has not been well-described (Dietert *et al.* 2000). Simple scaling assumptions do not adequately address variability in uptake via inhalation due to these age/developmental differences. Further, the collected data would provide a basis for assessing confidence in the description, thereby informing the magnitude of the interspecies and intra-human uncertainty/adjustment factors used in risk assessment.

This research would inform efforts to assign uncertainty factors for pharmacokinetic and pharmacodynamic differences (see below) and complement a Federal interagency collaborative effort to develop a suite of dosimetry models for oral, inhalation and dermal exposures. Such models have potential applications for both cancer and non-cancer endpoints. Development of such models therefore contributes to a harmonized approach for cancer and non-cancer risk assessment, and is

consistent with USEPA's *Methods for Derivation of Inhalation Reference Concentrations and Application of Inhalation Dosimetry* (USEPA 1994) and proposed carcinogen guidelines (USEPA 1996b).

E. Comparative Studies of Variability in Susceptibility to Toxic and Other Effects in Animals and People

Traditional protocols for animal toxicology experiments usually go to considerable lengths to minimize the variability among tested animals. This is because, in general, the more the variability, the larger the sample size required to demonstrate differences between the effects of experimental and control exposures. Variability is therefore deliberately restricted in most animal experimental work by using genetically homogenous animals, a single age class of subjects ("young adults" usually) and sometimes a single sex of animals. Before initiation of treatment the animals usually have been subject to relatively uniform environmental stimuli, including uniform and unchanging diet. Efforts are also made to maintain healthy animal colonies, as free of infections as possible. Thus the laboratory animal is not exposed to the diversity of living conditions of wilder, outbred populations. And, of course, there are also no deliberate neuroactive drug exposures (*e.g.*, narcotics, alcohol, tobacco, caffeine) unless they are explicit subjects of experimental study (Hattis 1996).

Studies are needed to experimentally assess in animal systems how much some or several of these common practices actually reduce variability in the doses producing defined toxic responses. Studies are also needed to assess the general distribution of comparative degrees of variability in toxic response for the free-living human population, relative to the types of animal groups usually used for toxicological testing. That is, how often is there a large difference in animal/human variability?

F. Procedures to Utilize Interindividual Variability Information in Cancer Risk Assessment

This would be an extension of the efforts described above under genetic determinants of variability (A), and (D) variability of mechanistic determinants of chemical disposition. Efforts are needed to elucidate an appropriate set of operational procedures to incorporate available generic and chemical-specific information on human interindividual variability in susceptibility to carcinogenesis into risk analyses utilized for risk management under a range of regulatory authorities and risk management criteria. For an exploration of some different potential implications of variability *vs* uncertainty for risk management under different regulatory authorities see Hattis and Anderson (1999), Krewski *et al.* (1999), Hattis and Minkovitz (1996), and Bois *et al.* (1996).

MODELS FOR INJURY RISK ASSESSMENT INCLUDING DEVELOPMENT AND CHARACTERIZATION OF RISK ESTIMATES AND EXPOSURE METRICS

James T. Wassell and John Bailer proposed the development of methodology for assessing risk and exposure metrics for occupational injury. The application of risk

assessment methods to this problem provides many opportunities for innovative development, but more importantly has the potential to significantly contribute to public health improvements. The incidence of occupational and non-occupational injuries is substantial, with a relatively early average age of occurrence for most serious injuries. In comparison to cancers and manifestations of cardiovascular disease, social effect measures such as years-of-potential-life-lost (YPLL) (Gilbert *et al.* 1998) are potentially greater for injury. The field of injury risk is just beginning to define relevant endpoints for social policy evaluation such as working lifetime risk (Fosbroke *et al.* 1997; See and Bailer 1998).

Among the most interesting methodological challenges is the definition of relevant causal and potentially confounding exposures. The most commonly available occupational injury data are based on the number of hours worked in the workplace and the size of the workforce. The proportion of time actually at risk (for example while using a particular type of machine or while engaged in material handling) is not distinguished from the portion of time worked while not at risk or at risk in other ways. Wassell (1989) discussed probability considerations for commonly used methods of statistical analysis of injury data. Poisson regression and other models of injury occurrence that would be more descriptive and helpful for evaluating the efficacy of interventions would include prospective or blind retrospective assessments. Modifiable targets for preventive action include factors such as the time spent operating unguarded machines, or the number of maintenance operations that could lead to a violation of lock-out/tagout precautions. Methods should account for personal confounding factors (*e.g.* worker age, job tenure, training) to control for non-engineering contributions to injury risks.

Another important area of research is the development of exposure-response models for chronic repetitive motion injuries, analogous to cancer multistage or classical toxicological probit dose response models. Such models should be based on considerations of physiological factors and data on the variability in the frequency and intensity of repetitive motion stresses created by particular work tasks, as well as variability in physiological responses to repeated subclinical injury events. They ideally should also be based on a mechanistic theory for how irreversible or very slowly reversible injury events happen. Finally, work is needed on priority-setting for injury prevention efforts.

ADAPTATIONS AND MODIFICATIONS TO EXISTING STANDARD PROCEDURES (*e.g.*, FOR DERIVING RFD'S)

There were three proposals aimed at improving on existing default procedures for risk assessment:

Further explorations of "data-derived" uncertainty/adjustment factors (UFs) through the development of

data bases for existing uncertainty factors

probabilistic distributions of UFs adapted to specific types of agents and effects

probabilistic approaches for reference dose (RfD) derivation

criteria for using data-derived factors

Development of quantitative guidance for expected risks posed by an RfD or, alternatively, the probability that the population threshold is below calculated RfD levels.

- Approaches to recognize influences such as age, latency, and pattern of exposure in bioassays and exposed human populations

The first two were originally developed by Lynne Haber, and the third was generated in group discussion. The proposals are briefly discussed in the following.

A. Data-Derived UFs

Traditionally, default UFs of 10-fold have been used for extrapolation from animals to humans, and for accounting for variability among humans due to sensitive populations and individuals. Development and examination of data bases, for example on human and rodent chronic toxicity for certain classes of agents, may provide the rationale for a different default for application in specific circumstances. A factor so identified is termed a "data derived UF;" a few assessments have been conducted using such factors (*e.g.*, Bogdanffy and Jarabek 1995; Bogdanffy *et al.* 1999; Dourson *et al.* 1998; IPCS 1998). Renwick (1991) and Renwick and Lazarus (1998) suggest interspecies and intraspecies UFs can each be broken into toxicokinetic and toxicodynamic components, based on the relative contributions of these components for a number of chemicals examined (*e.g.*, for the interspecies factor, a factor of 3.3 for each component). Similarly, reduction of the interspecies UF from 10 to 3 is used to address uncertainty in laboratory animal to human extrapolation for Reference Concentrations (RfCs) when dosimetric adjustments are made for species differences in toxicokinetics (USEPA 1994; Jarabek 1995a). As a further step, as outlined in the second bulleted section below, the use of distributions, rather than single point estimates, has been proposed (Hasseblad and Jarabek 1996; Baird *et al.* 1996; Hattis *et al.* 1999).

The issue of criteria for replacement of default UFs has been raised (*e.g.*, Meek 2000). (Since the workshop, the term "chemical-specific adjustment factors" has replaced the term "data derived UFs." Guidance for the use of data in the development of chemical-specific adjustment factors for interspecies differences and human variability have been developed by the IPCS, and are available at http://www.who.int/pcs/pubs/pub_list.htm.) Whether the critical determining factor for toxicity has been identified in developing the factors can be questioned. To derive a factor for the interspecies toxicokinetic component, laboratory animal to human ratios of the values of some human toxicokinetic parameters have been compiled, but it is usually not clear that these ratios are adequate surrogates for the ratios of critical tissue doses in the two species. Similarly, it is currently unclear how to translate information on human variability in key metabolic parameters into an uncertainty factor for human variability in kinetics. Descriptions of human variability in the critical pharmacokinetic parameters through the PBPK modeling of well-

characterized model chemicals acting via selected modes of action may help to elucidate whether such variability is adequately described by the variability in certain metabolic parameters (*e.g.*, a key enzyme's V_{max} or V_{max}/Km ratio).

Risk assessment applications involving uncertainty factors less that the traditional defaults of 10-fold for inter- and intraspecies variation are beginning to appear. The National Research Council (2000) has used such factors to establish acute exposure guideline levels (AEGLs) for highly hazardous substances such as mono and dimethylhydrazine, when the available data are deemed sufficient to support such practice. Further details of the risk assessment methodology used in establishing AEGLs are given by the National Research Council (2001).

B. Characterizing Uncertainty in RfDs and RfCs

Traditionally, RfDs and RfCs have been derived by dividing "no observed adverse effect levels" (NOAELs) observed in toxicological experiments or epidemiology by fixed uncertainty factors (although recently there has been some use of "benchmark doses" instead of NOAELs). A number of investigators have recently sought to replace fixed uncertainty factors with a probability density function, or PDF, that characterizes the uncertainty in the size of a "true but unknown" scaling factor (*e.g.*, Baird *et al.* 1996; Slob and Pieters 1997; Price *et al.* 1997; Swartout *et al.* 1998). For the UF of interest, the PDF is derived from data for that factor (*e.g.*, interspecies difference) for a reasonably large sample of chemicals. In addition to developing a PDF for the uncertainty factor, the uncertainty in the identification of a no effect level from animal study observations can be addressed (*e.g.*, Leisenring and Ryan 1992). Some residual risk is expected at the NOAEL due to the limited numbers of animals studied and other limitations of experimental design. The distributions of the UFs and for the animal no observed effect level may be convolved to develop a probabilistic distribution for the RfD. A probabilistic RfD has been developed for methylmercury (Clewell *et al.* 1999), based on a distribution of the input variables to the PBPK model used to convert hair mercury concentrations to a chronic intake rate.

PDFs for the RfD would provide useful information for both risk assessment and characterization, including uncertainty disclosure, information on the degree the RfD may be protective, and uncertainty in the overall health protection process. However, much work is needed before such distributions can be reliably derived, particularly the development of a statistical framework for the analysis. The derivations of PDFs for UFs typically involve comparisons of NOAELs from experiments conducted using different numbers of animals and experimental protocols. Proposed research includes the development of a statistical framework, for example to account for the uncertainties in the derived ratios, the residual risk at the NOAEL, and the updating of distributions based on chemical specific information. Work to date has not considered in detail how to take into account chemical-specific information in the development of the PDFs. The impact of ad hoc changes in UFs based on chemical specific data considerations on the characteristics of the appropriate PDF could be explored.

Dale Hattis raised a distinct but related research need to develop data and methodology to examine the assumption of true population thresholds embodied

in the traditional RfD concept. The recent work of Hattis et al. (1999) is based on the idea that human population distributions of susceptibility may be described by continuous lognormal or other more complex distributions, implying finite and potentially estimable risks at various levels of exposure above and below traditional RfDs. Also, when the disease of concern occurs in the population by the same mechanism as that of the toxicant in question, a population threshold is unlikely and some risk from exposure is expected. In this framework, RfDs would be defined as the human dose or exposure levels expected to produce no more than a specific incidence of harm at a minimal severity with a defined degree of confidence.

C. Influences of Age, Latency, and Pattern of Exposure and Other Factors

This is related to proposal #E discussed under the first program group above on variability. The work here would be to explore whether the traditional uncertainty factors adequately capture heterogeneity given the necessary limitations in toxicological study design.

CRITERIA FOR ACCEPTABILITY OF MECHANISTIC HYPOTHESES AND DATA

Rory Conolly proposed the initiation of a process to develop consensus criteria for the use on mechanistic information in risk assessment and management. He noted that the movement away from routine use of default assumptions towards more data-based approaches to risk assessment is in step with the increasing sophistication of laboratory methods for the study of biochemical mechanisms and the identification of trace levels of contaminants. These developments are raising questions of how data-based assessments should be structured and of the selection of endpoints for assessment.

Sensitive analytical techniques and the new biochemical techniques such as genomics and proteomics have the potential to link contamination at widespread environmental levels with changes in the expression levels of genes or the concentrations of specific proteins. In some cases, as for TCDD, biologically plausible hypotheses can be developed, suggesting that such early biochemical effects are the precursors of downstream frankly toxic effects (USEPA 2000). Dose response information for the early biochemical effects might be interpreted as indicative of the expected shape of the dose response curve for the downstream frankly toxic effect. As a result, risk assessors and risk managers are faced with several issues, including:

1. Do biochemical changes, such as a change in the expression level of a gene or the amount of a protein in a cell, constitute "adverse" effects?

2. How often and under what circumstances does the shape of the dose response curve for an early effect, such as a change in gene expression, inform us about the expected shape of the dose response curve for a frankly toxic effect further "downstream" in the causal pathway of harm?

3. Is homeostasis a determinant of the shape of the dose response curve? If so, what are the implications of homeostasis for the shape of the dose response curve?

4. Does the additivity to background argument (Crump *et al.* 1976) apply if homeostasis is operative or if pharmacokinetic nonlinearities exist?

5. When the understanding of mechanism is incomplete (as it always is), what criteria should be used to judge the acceptability of hypothetical linkages between early biochemical and later, frankly toxic effects?

6. How should mechanistic information be used in risk assessment when available data do not allow discrimination between alternative, plausible mechanistic hypotheses? Should a spectrum of assessments be developed based on the alternative hypotheses?

The broader risk assessment community, consisting not only of risk assessors but also of researchers collecting the critical data, need to carefully consider these questions in order to work toward shared understanding and consensus that is helpful for social decision-making. The critical needs are for (a) research design and development of mechanistic risk assessment models to be intelligently coordinated with each other and, (b) development of consensus criteria for the use of mechanistic data in risk assessment and risk management under different types of regulatory mandates. Workshops on the development of consensus criteria would be helpful in promoting reasoned discussion of these issues.

In a related proposal, "Extending the Range of Observation: Quantitative Relationships Between Key Biological Events to Aid Designation of Adversity and Identify Health Effects," Annie Jarabek emphasized the importance placed by the USEPA on characterizing the mode of action, defined as a chemical's influence on molecular, cellular or physiological functions (USEPA 1996; Jarabek 2000) in risk assessments. This requires a conceptual model that evaluates key events along the exposure-dose response continuum. Biomarkers data based in a mode-of-action framework essentially provide precursor lesion data and can serve as a basis for a parallelogram approach to extrapolation and determination of human homology for the health effect of interest (U.S. EPA 1994; Jarabek 1999). Thus, the framework provides for the extension of the range of observation, *e.g.*, for identification of biochemical or cellular events as measures of response, provided that causal links can be established to health outcome. She proposed research to evaluate the quantitative relationships among key events (*e.g.*, liver and cellular proliferation linked to tumor outcome) — from internal dose, to biologically effective dose, to various early effect indicators, to various outcome measures. This would provide a platform for the integration of diverse data, for example epidemiological data on effects in the population, and toxicological and mechanistic data acquired at the target tissue, cellular and subcellular levels. This work would also provide the necessary data to begin development of criteria for designation of adversity for use in risk assessment, *e.g.*, a specified degree of perturbation in cellular event such as 10% increase in cellular proliferation might be designated as a NOAEL or LOAEL. This work is important to accurately defining a given biomarker (defined by the NAS for exposure, effect, and susceptibility or combinations thereof) and distinguishing adaptive versus adverse effects.

MODELING ANALYSES FOR MULTIPLE ENDPOINT DATA

David Dunson proposed research to develop sophisticated statistical analytical tools to analyze multiple endpoint data, as is being generated in reproductive and neurodevelopmental studies. In recent years there has been increasing concern that exposure to chemicals with endocrine disrupting properties during development may have irreversible effects on reproductive, immune, and central nervous system function. This concern was formalized in the NAS report, *Pesticides in the Diets of Infants and Children*, which called for better information on the effects of pesticide exposure during development. In response to this report, testing designs have been implemented in which pregnant dams are dosed for the week before and after birth, and then the pups are dosed through puberty. Animals are tested at various points in the dosing period to ascertain effects on a variety of neurobehavioral, immunological, and reproductive outcomes.

Standard approaches for characterizing risk from toxicological studies are not ideal for multigenerational and developmental studies, in which multiple correlated endpoints are measured, as well as effects occurring across generations of related individuals. If the outcomes are considered separately and no adjustment is made for multiple comparisons, analyses will often detect some differences among dose groups (at, for example, $p = .05$) even if the chemical has no effect. However, standard adjustments for multiple comparisons make it very difficult to detect real effects, if present, because of the small numbers of animals tested and the large number of endpoints. An additional complication is that sick animals often die prior to being measured for outcomes that occur later in development. Such survival effects can produce biased estimates and misleading inferences. Another issue that arises in quantitative risk assessment is how to estimate a benchmark dose or virtually safe dose based on multiple correlated endpoints that are measured on different scales.

The purpose of this program would be to develop new approaches for assessing and characterizing risk in toxicology studies with multiple endpoints that are potentially measured on a variety of scales (*e.g.*, continuous, binary, ordinal). In particular, methods would be considered for reducing the dimensionality of the analysis, possibly by using a few "latent variables" underlying a set of measured outcomes (Dunson 2000). In addition, methods would be developed for estimating a dose level associated with a designated "acceptable" level of risk. One possibility would be to estimate the change with dose in the proportion of animals with a lower level of function than an average untreated animal (*e.g.*, by using a latent variable approach). The dose associated with a small change (*e.g.*, 1%) could be useful for policy makers deciding on permissible levels of exposure. A major objective would be to formulate a method that is readily interpretable by both toxicologists and risk managers.

Some important questions that need to be considered when developing this kind of method are

Should there be an adjustment for informative censoring from the deaths of sick animals before the end of the study?

2. Does the method have good operating characteristics in the small samples typical of these types of developmental toxicology studies?

3. Does the model account for correlation between different endpoints measuring a similar trait (*e.g.*, related neurobehavioral functions)?

4. Does the model account for correlation between endpoints measuring different traits (*e.g.*, unrelated neurobehavioral functions)?

5. Is the correlation structure realistic, but simple enough that the analysis yields information about all of the parameters, even in small samples?

6. Is there prior information that can be incorporated to improve the method? How can this information be combined with data from the current study?

7. Does the method account for different measurement scales for the different outcomes and for outcomes that do not follow standard distributions?

8. Can a latent variable or factor analytic type approach be used to simplify modeling and risk assessment (*e.g.*, by incorporating the dose effect relationship on one latent variable underlying the observed outcomes)?

9. If the latent variable approach is used, how many latent variable are there underlying the observed outcomes. How can the data be used to provide information about this choice?

10. Is the method robust (*e.g.*, to outliers, choice of parametric forms, etc.

11. Is the model identified by the data? Can reasonable constraints be added to ensure identifiability and improve efficiency?

12. Is the model easy enough to implement and interpret for non-statisticians to be used widely?

NEW MECHANISTIC MODELS OF CARCINOGENESIS

Two major topics were discussed under this heading:

- The implications for risk assessment modeling of genomic instability and webs of molecular pathological pathways to carcinogenesis (instead of a completely sequential set of k ordered stages in which particular genetic changes are acquired in a fixed order)

- The need to develop procedures (*e.g.* Markov chain Monte Carlo analysis) to represent the available biological information on signaling pathways and cellular responses mediating carcinogenesis)

Chao Chen highlighted some of the research opportunities related to this area

It is now widely accepted that cancer results from the accumulation of mutations and other genetic changes in the genes that directly control cell division, cell death, and differentiation (Bishop 1987; Weinberg 1989; Sugimura 1992; Williams *et al.*

1996). These include positive changes in growth/division signaling systems via "protooncogenes" and abrogation of specific growth/division control functions in tumor suppresser genes (Barrett 1993). For some human cancer sites, molecular biological tools have been used to elucidate the specific sequence of alterations in genetically determined functions both in full-fledged tumors and in tissue where the process is not yet complete (Fearon and Vogelstein 1990; Shi 1999; Mao *et al.* 1997; Grossman and Lefell 1997; Cordon-Cardo *et al.* 1994; Sekido *et al.* 1998).

Given these recent advances in molecular research, and increasing understanding of the different kinds of influences (*e.g.*, "nongenetic" events such as cell proliferation) on carcinogenesis that can be exerted by chemical exposures, risk assessors face major challenges in adapting risk assessment models to accommodate new mechanistic understanding. For example TCDD, acting in part through Ah receptors, triggers a variety of biological responses that can be divided into two broad categories: (1) metabolic changes associated with uptake and subsequent binding with other proteins, and (2) mitogenic processes manifested as DNA replication, cell division, and alterations in differentiation. Which of these are relevant in what ways quantitatively for predicting both cancer risks and a variety of non-cancer risks? Schwarz *et al.* (2000) recently reviewed several proposed possible influences on signaling pathways in relation to TCDD-induced antiapoptotic activities and concluded that many of the proposed pathways are not plausible. It would be highly misleading if a putative causal pathway were chosen for modeling that turned out not to be relevant to the endpoints of ultimate interest. As a general approach a model could be constructed that incorporates rate-limiting factors that are known to be in a relevant signaling pathway, and a black box could be used to represent unknown or uncertain steps. This type of situation is best handled with a newly developed type of mathematical tool that can be easily used to construct a stochastic model that incorporates all available biological information from activation of signaling pathways to cellular responses and tumor incidence (Tan and Chen 1998). Application of this kind of tool to risk assessment problems is a promising area for further exploration.

COMBINING DATA OF DIFFERENT TYPES IN RISK ANALYSES

This was a recurring theme emphasized by several participants in the course of the discussions. The main themes were to encourage

- Development of hierarchical Bayesian frameworks for cancer risk assessment to:

  —combine data from multiple sources of a variety of types (*e.g.*, data from different animal bioassays covering animals of different cancer sites, genders and species, and human epidemiology)

  —integrate data on the effects of a single chemical via different modes of action

  —develop probabilistic distributions of hazard potential

- Application of a control theory framework to

  —understand the influences of population heterogeneity in metabolic pathway activities on susceptibility and risk.

EXPLORING EVIDENCE AND MODELS FOR COMPLEX DOSE RESPONSE RELATIONSHIPS IN THE CONTEXT OF HOMEOSTASIS

Both Kenny Crump and Ralph Kodell submitted proposals for research on the toxicological phenomenon called hormesis, whereby a substance causes deleterious effects at high doses but a stimulatory response in the opposite direction at low doses (Calabrese *et al.* 1999). Dr. Crump's proposal was to develop and apply rigorous statistical tools to evaluate the pervasiveness of hormesis, while Dr. Kodell's proposal was to develop biologically plausible mathematical models that predict less-than-background risk of cancer at low doses. Under Dr. Crump's proposal, a toxicological data base comprised of data sets meeting a set of minimal criteria would be examined rigorously, controlling the false positive error rate while making the power to detect hormetic effects as large as possible. Meta-analytic procedures would be used to estimate the prevalence of hormesis in the data base, such as the procedure used by Crump *et al.* (1999) to estimate proportions of carcinogenic and anticarcinogenic chemicals in bioassays conducted under the National Toxicology Program. Under Dr. Kodell's proposal, biological data would be collected and biomathematical models of hormesis would be developed, to build on hypotheses (Andersen and Conolly 1998; Lutz 1998) and models (Bogen 1998; Lutz and Kopp-Schneider 1999) that have been proposed to support U-shaped or J-shaped dose response relationships for cancer. If, for example, toxic substances can affect homeostatic processes in such a way as to produce dose response relationships exhibiting less-than-background risk at low doses, then the default assumption that either genotoxicity or additivity to background will necessarily lead to low-dose linearity needs to be re-evaluated. Some others in the group are skeptical that a careful and rigorous analysis would lead to this result (Hattis 1997).

OVERALL CONCLUSION

The suggestions for research described above indicate the potential for greatly improved contributions from different fields (toxicology, molecular epidemiology, mathematical modeling, etc.) to improve dose response modeling. This can in turn improve the estimation of potential health risks and the consequences of different risk control choices. The field is the focal point of the dynamic interaction of data and theory relevant to important social policy concerns.

REFERENCES

Allen BC, Covington TR, and Clewell HJ. 1996. Investigation of the impact of pharmacokinetic variability and uncertainty on risks predicted with a pharmacokinetic model for chloroform. Toxicology 111:289-303

Andersen ME and Conolly RB. 1998. Mechanistic modeling of rodent liver tumor promotion at low levels of exposure: An example related to dose response relationships for 2,3,7,8-

tetrachlorodibenzo-p-dioxin. Human Experimental Toxicol 17, 683-90; discussion 701-4, 708-18

Bailer AJ and Dankovic DA. 1997. An introduction to the use of physiologically-based pharmacokinetc models in risk assessment. Stat Methods Medical Res 6:341-58

Bailer AJ, Stayner LT, Smith RJ, *et al.* 1997. Estimating benchmark concentrations and other non-cancer endpoints in epidemiology studies Risk Anal 17:771-80

Baird JS, Cohen JT, Graham JD, *et al.* 1996. Noncancer risk assessment: Probabilistic characterization of population threshold doses. Human Ecol Risk Assess 2(1):79-102

Barrett JC. 1993. Mechanisms of multistep carcinogenesis and carcinogen risk assessment. Environ Health Perspect 100:9-20

Bishop JM. 1987. The molecular genetics of cancer. Science 235:305-11

Bogdanffy MS and Jarabek AM 1995. Understanding mechanisms of inhaled toxicants: Implications for replacing default factors with chemical-specific data. Toxicol Lett 82/83:919-32

Bogdanffy MS, Sarangapani R, Plowchalk DR, *et al.* 1999. A biologically-based risk assessment for vinyl-acetate induced cancer and noncancer inhalation toxicity. Toxicol Sci 51:19-35

Bogen KT. 1998. Mechanistic model predicts a U-shaped relation of radon exposure to lung cancer risk reflected in combined occupational and US residential data. Human Experimental Toxicol 17:691-6

Bogen KT and Spear RC. 1987. Integrating uncertainty and interindividual variability in environmental risk assessment. Risk Anal 7:427-36

Bois FY. 1999. Analysis of PBPK models for risk characterization. Ann NY Acad Sci 895:317-37

Bois FY, Gelman A, Jiang J, *et al.* 1996. Population toxicokinetics of tetrachloroethylene. Arch Toxicol 70(6):347-55

Bois FY, Krowech G, and Zeise L. 1995. Modeling human interindividual variability in metabolism and risk: the example of 4-aminobiphenyl. Risk Anal 15:205-13

Bois FY, Zeise L, and Tozer TN. 1990. Precision and sensitivity of pharmacokinetic models for cancer risk assessment: Tetrachloroethylene in mice, rats, and humans. Toxicol Applied Pharmacol 102:300-15

Calabrese EJ, Baldwin LA, and Holland CD. 1999. Hormesis: a highly generalizable and reproducible phenomenon with important implications for risk assessment. Risk Anal 19:261-81

California Environmental Protection Agency. 2000. Draft Public Health Goal for Tetrachloroethylene in Drinking Water. Office of Environmental Health Hazard Assessment, Sacramento, CA, USA

Chemical Industry Institute of Toxicology (CIIT). September 28, 1999. Formaldehyde: Hazard Characterization and Dose Response Assessment for Carcinogenicity by the Route of Inhalation. Revised Edition. Chemical Industry Institute of Toxicology, Research Triangle Park, NC, USA

Chien JY, Thummel KE, and Slattery JT. 1997. Pharmacokinetic consequences of induction of CYP2E1 by ligand stabilization. Drug Metab Dispos 25:1165-75

Clewell HJ. 1995. The use of physiologically based pharmacokinetic modeling in risk assessment: A case study with methylene chloride. In: Olin S, Farland W, Park C, *et al.* (eds), Low-Dose Extrapolation of Cancer Risks: Issues and Perspectives. ILSI Press, Washington, DC, USA

Clewell HJ, Gearhart JM, Gentry PR, *et al.* 1999. Evaluation of the uncertainty in an oral Reference Dose for methylmercury due to interindividual variability in pharmacokinetics. Risk Anal 19:541-52

Clewell HJ and Jarnot BM. 1994. Incorporation of pharmacokinetics in non-carcinogenic risk assessment: Example with chloropentafluorobenzene. Risk Anal 14:265-76

Clewell HJ, III and Andersen ME. 1996. Use of physiologically-based pharmacokinetic modeling to investigate individual versus population risk. Toxicology 111:315-29

Cordon-Cardo C, Dalbagni G, Sarkis AS, *et al.* 1994. Genetic alterations associated with bladder cancer. Important Advances in Oncology 1994:71-83

Crump KS, Hoel DG, Langley CH, *et al.* 1976. Fundamental carcinogenic processes and their implication for low dose risk assessment. Cancer Res 36:2973-9

Crump KS, Krewski D, and VanLandingham C. 1999. Estimates of the proportion of chemicals that were carcinogenic or anticarcinogenic in bioassays conducted by the National Toxicology Program. Environ Health Perspect 107:83-8

Cullen AC and Frey HC. 1999. Probabilistic Techniques in Exposure Assessment—A Handblook for Dealing with Variability and Uncertainty in Models and Inputs. Plenum Press, NY, NY, USA

Dankovic D and Bailer AJ. 1994. The impact of exercise and intersubject variability on dose estimates for dichloromethane derived from a physiologically-based pharmacokinetic model. Fund Applied Toxicol 22:20-5

Dietert RR, Etzel RA, Chen D, *et al.* 2000. Workshop to identify critical windows of exposure to children's health: Immune and respiratory systems work group summary. Environ Health Perspect 108(suppl. 3):483-90

Dourson M, Maier A, Meek B, *et al.* 1998. Boron tolerable intake: Re-evaluation of toxicokinetics for data-derived uncertainty factors. Biol Trace Elem Res 66:453-63

Dunson DB. 2000. Bayesian latent variable models for clustered mixed outcomes. J R Stat Soc B 62 (part 2):355-66

Fearon ER and Vogelstein B. 1990. A genetic model for colorectal tumorigenesis. Cell 61:759-67

Finkel AM. 1987. Uncertainty, Variability and the Value of Information in Cancer Risk Assessmentt. Ph.D. Dissertation. Harvard University, Harvard School of Public Health, Cambridge, MA, USA

Finkel AM. 1995. A quantitative estimate of the variations in human susceptibility to cancer and its implications for risk management. In: Olin S, *et al.* (eds), Low-Dose Extrapolation of Cancer Risks: Issues and Perspectives, pp 297-328. ILSI Press, Washington, DC, USA

Fosbroke DE, Kisner SM, and Myers JR. 1997. Working lifetime risk of occupational fatal injury. Am J Ind Med 31:459-67

Gilbert SJ, Bailer AJ, Stayner LT. 1998. Years of Potential Life Lost due to Occupational Fatal Injury in the United States. Human Ecol Risk Assess 4(6):1321-35

Grossman D and Leffell DJ. 1997. The molecular basis of nonmelanoma skin cancer: new understanding. Archives Dermatol 133(10):1263-70

Hasselblad V and Jarabek AM. 1996. Dose-response analysis of toxic chemicals. In: Berry DA and Stangle DK (eds), Bayesian Biostatistics. Marcel Dekker, Inc., NY, NY, USA

Hattis D. 1996. Human interindividual variability in susceptibility to toxic effects–from annoying detail to a central determinant of risk. Toxicology 111:5-14

Hattis D. 1997. Invited comment on Heitzman, M., and Wilson, R. low dose linearity: The rule or the exception. Belle Newsletter 6:21-4

Hattis D. 1998. Strategies for assessing human variability in susceptibility, and variability to infer human risks. In: Neumann DA and Kimmel CA (eds) Human Variability in Response to Chemical Exposure: Measures, Modeling, and Risk Assessment, pp 27-57. CRC Press, Boca Raton, FL, USA

Hattis D and Anderson E. 1999. What should be the implications of uncertainty, variability, and inherent 'biases'/'conservatism' for risk management decision making? Risk Anal 19:95-107

Hattis D, Banati P, Goble R, *et al.* 1999. Human interindividual variability in parameters related to health risks. Risk Anal 19:705-20

Hattis D and Barlow K. 1996. Human Interindividual Variability in Cancer Risks—Technical And Management Challenge. Health Ecol Risk Assess 2:194-220

Hattis D and Burmaster DE. 1994. Assessment of variability and uncertainty distributions for practical risk analyses. Risk Anal 14:713-30

Hattis D and Minkowitz WS. 1996. Risk evaluation: criteria arising from legal traditions and experience with quantitative risk assessment in the United States. Environmental Toxicol Pharmacol 2:103-9

Hattis D and Silver K. 1994. Human interindividual variability—a major source of uncertainty in assessing risks for non-cancer health effects. Risk Anal 14:421-31

Integrated Risk Information System (IRIS). October 1999. Toxicological review of ethylene glycol monobutyl ether (EGBE) CAS no. 111-76-2. National Center for Environmental Assessment, U.S. Environmental Protection Agency, Washington, DC, USA

Integrated Risk Information System (IRIS). May 2000. Toxicological review of vinyl chloride CAS no. 75-01-4. National Center for Environmental Assessment, U.S. Environmental Protection Agency, Washington, DC, USA

Internation Programme for Chemical Safety (IPCS). 1998. Boron. Environmental Health Criteria. 204. World Health Organization, Geneva, Switzerland

Jarabek AM. 1995a. Interspecies extrapolation based on mechanistic determinants of chemical disposition. Human Ecol Risk Assess 1(5):641-62

Jarabek AM. 1995b. The application of dosimetry models to identify key processes and parameters for default dose response assessment approaches. Toxicol Lett 79:171-84

Jarabek AM. 1999. Application requirements for biomarkers in risk assessment. Presented at Biomarkers: Taking stock – An EPA/NIEHS In-house Workshop on applying biomarker research, 30-31 August 1999, Research Triangle Park, NC, USA (Manuscript in preparation)

Jarabek AM. 2000. Mode of Action: Framework for dosimetry model development. Presented in: Mode-of-Action Dosimetry: An interagency project to develop models for inhalation, oral and dermal disposition. Annual meeting of the Society for Risk Analysis, 4 – 6 December, Arlington, VA, USA (Manuscript in preparation)

Jonsson F and Johanson G. 2000. Improving the reliability of parameters in PBTK models using Markov Chain Monte Carlo simulation. (Abstract No. 422) Toxicologist 54:90

Kohn MC. 1995. Achieving credibility in risk assessment models. Toxicol Letters 79:107-14

Kohn MC. 1997. The importance of anatomical realism for validation of physiological models of disposition of inhaled toxicants. Toxicol Appl Pharmacol 147:448-58

Krewski D, Rai SN, Zielinski JM, *et al.* 1999. Characterization of uncertainty and variability in residential radon cancer risks. Annals NY Academy Sciences 895:245-72

Leisenring W and Ryan L. 1992. Statistical properties of the NOAEL. Reg Toxicol Pharmacol 15(2):161-71

Lutz WK. 1998. Dose response relationships in chemical carcinogenesis: superposition of different mechanisms of action, resulting in linear-nonlinear curves, practical thresholds, J-shapes. Mutation Res 405:117-24

Lutz WK and Kopp-Schneider. 1999. Threshold dose response for tumor induction by genotoxic carcinogens modeled via cell-cycle delay. Toxicol Sciences 49:110-5

Manton KG and Stallard E. 1979. Maximum likelihood estimation of a stochastic compartment model of cancer latency: lung cancer mortality among white females in the U.S. Computers and Biomedical Res 12:313-25

Manton KG, Stallard E, and Vaupel J, 1986. Alternative models for the heterogeneity of mortality risks among the aged. J Am Stat Assoc 81:635-44

Mao L, Lee JS, Kurie JM, *et al.* 1997. Clonal genetic alterations in the lungs of current and former smokers. J National Cancer Institute 89(12):857-62

Meek ME, Ohanian E, Renwick A, *et al.* 2000. Guidelines for application of data-derived uncertainty factors in risk assessment. Submitted.

National Research Council. 2000. Acute Exposure Guideline Levels for Selected Airborne Chemicals, Volume 1. National Academy Press, Washington, DC, USA

National Research Council. 2001. Standing Operating Procedures for Developing Acute Exposure Guideline Levels for Hazardous Substances. National Academy Press, Washington, DC (*in press*).

Occupational Safety and Health Administration (OSHA). 1997. Occupational exposure to methylene chloride; final rule. Fed Reg 62(7):1493-619

Price PS, Keenan RE, Swartout JC, *et al.* 1997. An approach for modeling noncancer dose responses with an emphasis on uncertainty. Risk Anal 17(4):427-37

Renwick AG. 1991. Safety factors and establishment of acceptable daily intakes. Food Add Contam 8(2):135-50

Renwick AG and Lazarus NR. 1998. Human variability and noncancer risk assessment—an analysis of the default uncertainty factor. Regul Toxicol Pharmacol 27(1 Pt 1):3-20

Schlosser PM and Bogdanffy MS. 1999. Determining modes of action for biologically based risk assessments. Regul Toxicol Pharmacol 30:75-9

Schwarz M, Buchmann A, Stinchcombe S, *et al.* 2000. Ah receptor ligands and tumor promotion: survival of neoplastic cells. Toxicol Letters 112-113:69-77

See K and Bailer AJ. 1998. Estimates of lifetime risk of occupational fatal injury from age-specific rates. Human Ecol Risk Assess 4(6):1309-19

Sekido Y, Fong KM, and Minna JD. 1998. Progress in understanding the molecular pathogenesis of human lung cancer. Biochimica et Biophysica Acta 1378:F21-59

Shi ST, Yang FY, Wang LD, *et al.* 1999. Role of p53 gene mutations in human esophageal carcinogenesis; results from immunohistochemical and mutation analyses of carcinomas and nearby non-cancerous lesions. Carcinogenensis 20:591-7

Slob W and Pieters MN. 1997. A Probabilistic Approach for Deriving Acceptable Human Intake Limits and Human Health Risks from Toxicological Studies: General Framework. 620110005. Rijksinstituut voor Volksgezondheid en Milieu, National Institute of Public Health and the Environment, The Netherlands

Sugimura T. 1992. Multistep carcinogenesis: A 1992 perspective. Science 258:603-8

Swartout JC, Price PS, Doruson ML, *et al.* 1998. A probabilistic framework for the reference dose. Risk Anal 18(3):271-82

Tan WY and Chen C. 1998. Stochastic modeling of carcinogenesis: Some new insights. Mathematical Computer Modeling 28(11):49-71

Thompson KM. 1999. Developing univariate distributions from data for risk analysis. Human and Ecol Risk Assess 5:755-83

USEPA (U.S. Environmental Protection Agency). October1994. Methods for Derivation of Inhalation Reference Concentrations and Application of Inhalation Dosimetry. EPA-600/8-90-066F. Office of Health and Environmental Assessment, Environmental Criteria and Assessment Office, Research Triangle Park, NC, USA

USEPA (U.S. Environmental Protection Agency). 1996a. Air Quality Criteria for Particulate Matter – Volume II. EPA-600/P-95-001bF. National Center for Environmental Assessment, Research Triangle Park, NC. Available from NTIS, Springfield, VA, USA; PB96-168224.

USEPA (U.S. Environmental Protection Agency). 1996b. Proposed guidelines for carcinogen risk assessment. Federal Register 61(79):17960-8011

USEPA (U.S. Environmental Protection Agency). 2000. Exposure and Human Health Reassessment of 2,3,7,8-Tetrachlorodibenzo-p-Dioxin (TCDD) and Related Compounds. Part II: Health Assessment for 2,3,7,8-Tetrachlorodibenzo-p-dioxin (TCDD) and Related Compounds. Chapter 8. Dose response Modeling for 2,3,7,8-TCDD. NCEA-I-0835, May 2000, SAB Review Draft.

Wassell JT. 1998. Probability Models of Occupational Injury. Human Ecol Risk Assess 4(6):1275-83

Wienberg RA. 1989. Ocogenes, antioncogenes, and the molecular bases of multistep carcinogenesis. Cancer Res 49:3713-21

Williams JR, Russell J, Dicello JF, *et al.* 1996. The genotype of the human cancer cell: Implications for risk analysis. Mutation Res 365:17-42