



## Advancing human health risk assessment: Integrating recent advisory committee recommendations

Michael Dourson, Richard A. Becker, Lynne T. Haber, Lynn H. Pottenger,  
Tiffany Bredfeldt & Penelope A. Fenner-Crisp

To cite this article: Michael Dourson, Richard A. Becker, Lynne T. Haber, Lynn H. Pottenger, Tiffany Bredfeldt & Penelope A. Fenner-Crisp (2013) Advancing human health risk assessment: Integrating recent advisory committee recommendations, Critical Reviews in Toxicology, 43:6, 467-492, DOI: [10.3109/10408444.2013.807223](https://doi.org/10.3109/10408444.2013.807223)

To link to this article: <https://doi.org/10.3109/10408444.2013.807223>



© 2013 The Author(s). Published by Taylor & Francis.



Published online: 12 Jul 2013.



Submit your article to this journal [↗](#)



Article views: 2345



View related articles [↗](#)



Citing articles: 7 View citing articles [↗](#)

## REVIEW

## Advancing human health risk assessment: Integrating recent advisory committee recommendations

Michael Dourson<sup>1</sup>, Richard A. Becker<sup>2</sup>, Lynne T. Haber<sup>1</sup>, Lynn H. Pottenger<sup>3</sup>, Tiffany Bredfeldt<sup>4</sup>, and Penelope A. Fenner-Crisp<sup>5</sup><sup>1</sup>Toxicology Excellence for Risk Assessment, Cincinnati, OH, USA, <sup>2</sup>American Chemistry Council, Washington, DC, USA, <sup>3</sup>The Dow Chemical Company, Midland, MI, USA, <sup>4</sup>Texas Commission on Environmental Quality, Austin, TX, USA, and <sup>5</sup>Independent Consultant, North Garden, VA (USEPA retired), USA

## Abstract

Over the last dozen years, many national and international expert groups have considered specific improvements to risk assessment. Many of their stated recommendations are mutually supportive, but others appear conflicting, at least in an initial assessment. This review identifies areas of consensus and difference and recommends a practical, biology-centric course forward, which includes: (1) incorporating a clear problem formulation at the outset of the assessment with a level of complexity that is appropriate for informing the relevant risk management decision; (2) using toxicokinetics and toxicodynamic information to develop Chemical Specific Adjustment Factors (CSAF); (3) using mode of action (MOA) information and an understanding of the relevant biology as the key, central organizing principle for the risk assessment; (4) integrating MOA information into dose–response assessments using existing guidelines for non-cancer and cancer assessments; (5) using a tiered, iterative approach developed by the World Health Organization/International Programme on Chemical Safety (WHO/IPCS) as a scientifically robust, fit-for-purpose approach for risk assessment of combined exposures (chemical mixtures); and (6) applying all of this knowledge to enable interpretation of human biomonitoring data in a risk context. While scientifically based defaults will remain important and useful when data on CSAF or MOA to refine an assessment are absent or insufficient, assessments should always strive to use these data. The use of available 21st century knowledge of biological processes, clinical findings, chemical interactions, and dose–response at the molecular, cellular, organ and organism levels will minimize the need for extrapolation and reliance on default approaches.

## Keywords

Biomonitoring equivalents, chemical specific adjustment factor (CSAF), combined exposures, critical effect, dose-dependent transitions, harmonized dose–response assessment, mode of action (MOA), problem formulation, safe dose

## History

Received 28 December 2012  
Revised 16 May 2013  
Accepted 17 May 2013  
Published online 8 July 2013

## Table of Contents

Abstract	467
Introduction	467
Selected committee deliberations	468
Problem formulation linked to risk management solution	468
Evolution of The “Safe” Dose and Its Related Safety Factor(s)	471
From critical effects to mode of action (MOA)	473
Low-dose extrapolation: transition from defaults to mode of action	
(MOA) understanding	477
Underlying assumptions	477
Using mode of action (MOA) information	478
Dose–response modeling	479
Areas of future growth in understanding MOA	481
Cumulative risk and mixtures	481
Biomonitoring	484
Discussion	486
Acknowledgements	487
Declaration of interest	487
References	487

## Introduction

Since the time of the seminal publication on human health risk assessment by the U.S. National Academy of Sciences/National Research Council (NRC, 1983), the risk assessment landscape, once bereft of methods, insights, and research, is now brimming—to the point where expert committees are needed to sort through the resultant cornucopia. Many of these committee deliberations are carefully and thoughtfully drafted, being based on a well-planned series of research efforts and workshops. Other deliberations are more limited, with some useful insights gained at the expense of other ideas in need of better articulation or reconsideration. Specifically with regards to dose–response assessment, one of the component parts of the NRC (1983) paradigm, some committee recommendations have been mutually supportive, while others are clearly contradictory. The contradictory recommendations sometimes lead to differences in the development of risk assessment methods, often lead to differing interpretations of risk information, and occasionally result in remarkably different risk management decisions for the same chemical substance.

Understanding the biological basis of disease and its profile in humans is foremost in any medical discipline, with toxicology and epidemiology being no exception. In cases where the medical science lies more with prevention, such as with these latter two disciplines, understanding the etiology of any disease is often more difficult due in part to multiple potential origins and the interplay of risk factors. For example, in evaluating the significance of body weight loss in a 2-year study, where the chemical is in the food and the experimental animal can eat as much as it wants, a risk assessor should consider this loss as adverse only in relationship to the health of control animals, since often, the controls will overeat and not be as healthy as the experimental animals. Similarly, low-dose extrapolation of epidemiology data should consider the underlying biology and information on the presence or absence of precursor endpoints in the dose range of interest and other available Mode of Action (MOA) information, and not rely on linear regressions without prejudice. The guidance documents and committee reports discussed in this article provide perspectives on how to incorporate biological information on normal physiology and disease mechanisms to interpret toxicological and epidemiologic information.

Evolving technologies, such as those suggested by the NRC report for Toxicity Testing in the 21st Century (NRC, 2007a), can also help elucidate the biological basis of disease and inform the assessment of response in sensitive humans at low doses. The current defaults that toxicologists and epidemiologists often use for their dose–response assessments should not constrain the use of the full extent of this new technology. Likewise, risk assessment theory has similarly evolved. Specifically, risk assessment scientists now routinely promote the following: (1) development of a problem formulation (PF) step prior to the assessment to focus effort and resources, (2) use of chemical-specific adjustment factors (CSAFs) from empirical data rather than default uncertainty factors, (3) consideration of MOA information early in the assessment process, and (4) evaluation of dose–response assessment with human relevance (HR) frameworks. These evolved concepts have been developed by a number of national, international, and multinational scientific bodies, and encouraged by the NRC (2007a, 2009) and many others, such as the Alliance for Risk Assessment (ARA, 2013). They now form the basis of risk assessment work worldwide, and are the standards against which new assessments should be judged.

These four concepts will also serve as an integrating structure for this discourse, which will address areas of consistency and areas of conflict among the various committee and agency recommendations.

As in any scientific review, it is important to specify what topics will not be covered. In this review, we will not discuss in any depth, screening level dose–response assessment (other than Hazard Index (HI)), exposure assessment, risk characterization, or risk communication, despite the importance of these topics. Nor will we focus on radiation standards of the National Ambient Air Quality Standards (NAAQS) of the US EPA. In the case of the radiation standards, the latest guidance document from the Committee on Biological Effects of Ionizing Radiation VII (BEIR, 2006) is available. In the case

of the NAAQS, Bachmann (2007) summarizes the history of setting NAAQS, and McClellan (2011) emphasizes the role of scientific information in informing the EPA Administrator's policy judgments on the level and statistical form of the NAAQS for a particular indicator and averaging time for a specific criteria pollutant.

Rather, we will focus on hazard identification and dose–response assessment, including the dichotomy of the practice in some organizations of separate default cancer and non-cancer extrapolations, and differing approaches to protecting sensitive individuals. Concordant recommendations among various committees will be highlighted; conflicting recommendations will be resolved, if possible, on the biological basis of adverse effect and through an understanding of the underlying PF/CSAF/MOA/HR frameworks.

## Selected committee deliberations

### Problem formulation linked to risk management solution

The concept of including problem formulation and a planning and scoping exercise prior to beginning the analysis phase of a risk assessment is generally embraced positively by all parties engaged in or affected by risk assessment or risk management decisions. Many parties, both outside and inside of the government (particularly at the U.S. Environmental Protection Agency; US EPA) have presented visions of how these pre-assessment elements would be incorporated, in principle, into the process. These visions are remarkably consistent with one another (see US EPA, 1992, 1998, 2000, 2006a, 2007; NRC, 1993, 1994, 1996, 2008a, 2009). The authors, however, have seen a significant level of concern expressed by parties outside of the agency that US EPA is only paying lip service to its purported commitment to implementing problem formulation and planning and scoping into its risk assessment/risk management process. In contrast to this perception by some, we assert that the US EPA routinely includes problem formulation, planning and scoping in its risk assessment and management work, as described in the remainder of this section.

In the first of an ever-growing series of publications from the NRC, the authors of the 1983 NRC report observed that risk assessments and related regulatory decisions issued by federal agencies have been “bitterly controversial.” Among the Committee's key recommendations was “that regulatory agencies take steps to establish and maintain a clear *conceptual* [emphasis added] distinction between assessment of risks and consideration of risk management alternatives; that is, the scientific findings and policy judgments embodied in risk assessments should be explicitly distinguished from the political, economic, and technical considerations that influence the design and choice of regulatory strategies.”

Since then, risk assessments and related regulatory decisions issued by federal agencies have continued to be the subject of heated criticism. Among the aspects criticized is an ongoing and apparent dissonance between the construct and content of the hazard/risk assessment and the construct of the regulatory decision. In US EPA's experience, this criticism has been leveled both from within the agency and from many outside sources, including the affected stakeholders. As a

1994 NRC report noted “Several commenters have concluded that the conceptual separation of risk assessment and risk has resulted in procedural separation to the detriment of the process.”

Based in part on this series of NRC reports, the US EPA began using the concept of problem formulation about twenty years ago, with the goal of helping to provide risk assessments that better fit the decision-makers’ needs (US EPA, 1992; NRC, 1993). The USEPA’s framework for ecological risk assessment, later incorporated into the agency’s 1998 ecological risk assessment guidelines, described an initial phase, to occur before any effort is expended on the risk assessment itself, as problem formulation.

Problem formulation includes a preliminary characterization of exposure and effects, as well as examination of scientific data and data needs, policy and regulatory issues, and site-specific factors to define the feasibility, scope, and objectives for the ecological risk assessment. The level of detail and the information that will be needed to complete the assessment also are determined (US EPA, 1992).

This phase was meant to include a planning discussion between the risk assessor(s) and the risk manager(s), not for the risk manager to provide the expected “answer” but, rather, to clarify expectations by laying out for all participants information such as what is already known, what data need to be developed and the context in which this information would be used. Importantly, these guidelines acknowledge that “interested parties,” in addition to the agency’s risk assessors and risk managers, may “take an active role in planning, particularly in goal development.” The guidelines describe interested parties, also called “stakeholders,” as:

Federal, State, tribal, and municipal governments, industrial leaders, environmental groups, small-business owners, landowners, and other segments of society concerned about an environmental issue at hand or attempting to influence risk management decisions. Their involvement, particularly during management goal development, may be key to successful implementation of management plans since implementation is more likely to occur when backed by consensus. Local knowledge, particularly in rural communities, and traditional knowledge of native peoples can provide valuable insights about ecological characteristics of a place, past conditions, and current changes. This knowledge should be considered when assessing available information during problem formulation (USEPA, 1998).

Within US EPA, only the Office of Pesticide Programs retains, with rare exception, both the risk assessment and risk management functions related to its legislative mandates (as per PF-C and MD). The other offices whose regulatory responsibilities depend, in part, on risk assessment, have yielded some, if not all, of their assessment tasks to a separate office. It could be said that this “solution” actually has impeded the agency from implementing its own problem formulation/planning and scoping framework(s) in many

specific instances, because of the absence of adequate collaboration and coordination between the risk assessors and the risk managers.

As noted above, although the US EPA had embraced formulation as the first step in developing a risk assessment, a series of NRC reports over the last two decades appear to express the opinion that problem formulation is only infrequently practiced by the US EPA and others conducting risk assessments. While this criticism might have been warranted at the time the 1994 and 1996 NRC reports were developed, it was misguided by the time the 2009 NRC report was underway. The existence of several generic guidance documents and many existing examples of their application (detailed below) seems to have been missed or ignored.

Improved planning and attention to the uses of the risk assessment were recommended by the NRC committee studying the US EPA’s implementation of the 1990 Clean Air Act amendments (NRC, 1994); it stated that such planning will aid in efficient resource allocation. That committee recommended that “the ‘Red Book’ paradigm should be supplemented by applying a cross-cutting approach that addresses the following six themes in the planning and analysis phases: default options, validation, data needs, uncertainty, variability, and aggregation.” Finally, the Committee expressed support for implementation of a tiered, iterative risk assessment approach.

The importance of problem formulation in the early stages of a risk assessment, and incorporation of an iterative process with feedback was further emphasized in the 1996 NRC report. In addition, the Presidential/Congressional Commission on Risk Assessment and Risk Management (1997) emphasized the importance of this initial step in designing a risk assessment, stating, “The problem/context stage is the most important step in the [Commission’s] Risk Management Framework.” Both the NRC and Presidential/Congressional Commission committees noted the importance of including all affected parties in the discussion, early and often, rather than restricting the discussion solely to agency risk assessors and risk managers. This does not necessarily mean that these affected parties will have a seat at the table when the final assessment or regulatory decision is made, but, rather, that they have had an opportunity to provide information that may help to make the assessment and associated decision(s) more complete and robust. Particularly good examples of substantive stakeholder involvement in planning and executing risk assessment and regulatory decisions can be seen in the processes employed by US EPA’s Office of Solid Waste and Emergency Response as its regional offices develop site-specific assessments (US EPA, 1997, 1999, 2001) and by the Office of Pesticide Programs as it implements the 1996 Food Quality Protection Act (US EPA, 2011a, 2011b, 2011c).

The 2009 NRC report focuses a great deal of attention on the design of risk assessments, devoting an entire chapter to this topic. It includes a schematic described as a “framework for risk-based decision-making that maximizes the utility of risk assessment.” Inferred to be a novel approach to this issue, the NRC framework looks remarkably like the framework schematics included in many of USEPA’s

already-published guidance documents (e.g. US EPA, 1992, 1998, 2000, 2001, 2003a, 2006a, 2007). Each of these frameworks usually includes three general phases, the first presenting concepts of problem formulation, planning and scoping, the second reflecting the risk assessment phase and, the third focused on the integration of other relevant factors (e.g. economics, technology, political considerations) to reach and communicate the management decision(s). The NRC (2009) Committee noted that the conceptual framework is missing from other agency guidance, although it is unclear to what “other guidance” they were referring. The NRC framework, however, does incorporate a level of detail not seen in most of USEPA’s framework documents, including specific questions in each of the three phases (Phase I: Problem formulation and scoping; Phase II: Planning and conduct of the risk assessment; Phase III: Risk Management).

Furthermore, the NRC Committee was very clear that it saw value in crafting a risk assessment that “ensures that its level and complexity are consistent with the needs to inform decision-making.” The 2009 NRC framework also reinforces the importance of having “formal provisions for internal and external stakeholder involvement at all stages.” The Committee also recommended that USEPA pay increased attention to the design of risk assessment in its formative stages and that USEPA adopt a framework for risk-based decision-making that embeds the Red Book risk assessment paradigm into a process with (1) initial problem formulation and scoping, (2) upfront identification of risk-management options, and (3) use of risk assessment to discriminate among these options.

Unfortunately, these recommendations do not necessarily mean that the NRC framework is better than existing ones, including those of US EPA. In fact, the agency is often asking the same questions when it implements its frameworks for specific cases, but one needs to read and study the specific case to understand its application.

Furthermore, although problem formulation was initially addressed at US EPA in the context of ecological risk assessment, a number of agency-wide and/or Office of Research and Development guidance documents that include an analysis phase for both ecological and human health risk assessment now incorporate the concept of problem formulation as the critical first step in the risk assessment process. Some examples of generic guidance include the Risk Characterization Handbook (US EPA, 2000), the Framework for Cumulative Risk Assessment (US EPA, 2003a), the Framework for Assessing Health Risks of Environmental Exposures to Children (US EPA, 2006a) and the Framework for Metals Risk Assessment (US EPA, 2007). The Risk Characterization Handbook contains several case studies of both human health and ecological concerns, each of which includes a discussion of how problem formulation was implemented. The Framework for Assessing Health Risks of Environmental Exposures to Children was developed as the result of a collaborative effort with the International Life Sciences Institute (ILSI), which sponsored a multi-stakeholder, multi-disciplinary workshop to craft the framework (Daston et al., 2004; Olin & Sonawane, 2003).

Moreover, most US EPA program offices and regions also have crafted a set of principles tailored to their specific

circumstances (e.g. US EPA 1999, 2001, 2011d). Examples include:

- The Office of Pesticide Program’s (OPP’s) Pesticides Registration Review Process, implemented after completion of the Food Quality Protection Act-mandated tolerance reassessment (US EPA, 2006b); currently there are dockets open for 240 registered active ingredients undergoing reevaluation of their regulatory status (US EPA, 2012b);
- The process of the Office of Air Quality Planning and Standards (OAQPS) for reviewing the National Ambient Air Quality Standards (NAAQS; US EPA, 2009); this process is currently being used in the reassessment of lead (US EPA 2011e) and the oxides of nitrogen (US EPA, 2012c);
- The Office of Water’s (OW’s) draft framework for integrated municipal and wastewater plans of its National Pollutant Discharge Elimination System (NPDES) program (US EPA, 2012d); and
- The Multi-criteria Integrated Resource Assessment (MIRA) approach employed by Region III (US EPA, 2003b); specific examples of its application are listed on the Region’s MIRA website (<http://www.epa.gov/reg3esd1/data/mira.htm>).

The concept of problem formulation also has been embraced internationally through the leadership of the World Health Organization (WHO), especially its International Programme on Chemical Safety (IPCS), with significant involvement from US EPA. Recent publications that acknowledge problem formulation as a critical component of the risk assessment/risk management paradigm include:

- Integrated Risk Assessment (Birnbaum et al., 2001; Suter et al., 2003);
- Environmental Health Criteria 237- Principles for Evaluating Health Risks in Children Associated with Exposure to Chemicals (WHO IPCS, 2006);
- Uncertainty and Data Quality in Exposure Assessment. Part 1. Guidance Document on Characterizing and Communicating Uncertainty in Exposure Assessment, Harmonization Project Document No. 6 (WHO IPCS, 2008);
- Environmental Health Criteria 239 - Principles for Modeling Dose-Response for the Risk Assessment of Chemicals (WHO IPCS, 2009a);
- Environmental Health Criteria 240 - Principles and Methods for the Risk Assessment of Chemicals in Food (WHO IPCS, 2009b; Renwick et al, 2003);
- Characterization and Application of Physiologically Based Pharmacokinetic (PBPK) Models in Risk Assessment. (WHO IPSC, 2010);
- Risk Assessment of Combined Exposure to Multiple Chemicals: A WHO/IPCS Framework (Meek et al., 2011);
- Guidelines for Drinking-water Quality-Fourth Edition (WHO, 2011).
- Microbial Risk Assessment Guideline Pathogenic Microorganisms With Focus on Food and Water (USDA, 2012).

Expert groups and world health organizations have nearly always used a problem formulation construct in the

deliberations of their assessment work, but this construct has not always been apparent or consistent.

Recommendations that have emerged from this analysis and related efforts are:

- (1) The concept of problem formulation as a prelude to a risk assessment work is generally, and should be uniformly, embraced globally by all health organizations.
- (2) Differences in risk management decisions, and in the products of the individual components of hazard characterization, dose–response assessment, exposure assessment, and risk characterizations, should be expected based on different problem formulations.
- (3) Risk management input on problem formulation, with its associated planning and scoping, is essential in order for risk assessment scientists to develop useful information. This upfront identification of risk management options should not be seen as changing or subverting the scientific process of risk assessment.

### Evolution of The “Safe” Dose and Its Related Safety Factor(s)

The concept of a safe dose is based upon the identification of a threshold<sup>1</sup> for an adverse effect.<sup>2</sup> This threshold is based on an experimentally determined Lowest Observed Adverse Effect Level (LOAEL), and its matching experimentally determined subthreshold dose, the No Observed Adverse Effect Level (NOAEL), the latter of which is adjusted to the safe dose through the use of a composite safety factor that is determined based on the available data. This concept has been in use since the late 1950s to establish safe dose in order to protect public health from potential chemical exposures. Exceedances of these safe doses have been used to describe situations of potential risk associated with such exposures to the public. This concept was built on two major assumptions: that protecting against the critical effect<sup>3</sup> protects against subsequent adverse effects, and that the use of a safety factor (now commonly referred to as uncertainty factor) lowers the acceptable exposure level to a resultant “safe” dose, that is, one below the range of the possible thresholds of the critical effect in humans, including sensitive subgroups.

This safe dose was called the Acceptable Daily Intake (ADI) and was used for oral exposure to chemical contaminants and approved food additives. Several historical accounts describe early deliberations on this concept (e.g. Clegg, 1978; Dourson & DeRosa, 1991; Kroes et al., 1993; Lu, 1988; Truhaut, 1991; Zielhuis & van der Kreek, 1979).

Although quite useful, a general problem with this concept has been that its key features, that is, the element of judgment required to define a NOAEL, and determination of an appropriate safety factor based upon the content and quality of the underlying database, did not allow a ready incorporation of dose–response data to refine the estimate. Starting after the 1970s, several initially separate series of research efforts or deliberations occurred that prompted the evolution of the safe dose and related safety factor concept.

The first effort started with Zielhuis & van der Kreek (1979) who investigated the use of safety factors in the occupational setting. Similar to these investigators, the US EPA separately reviewed oral toxicity data for human sensitivity, experimental animals to human extrapolation, insufficient study length (e.g. 90-day study only), and absence of dose levels without adverse effects (Dourson & Stara, 1983). Typically, the use of all of these factors would occur during the derivation of a “safe dose” for data-poor chemicals. Afterwards, in light of the then-recent NRC (1983) publication, US EPA changed its parlance to better reflect a separation of risk assessment and risk management. “Safety factor” became “uncertainty factor” and “ADI” became “Reference Dose<sup>4</sup> (RfD)” (Barnes & Dourson, 1988). Other organizations (e.g. U.S. Food and Drug Administration, WHO/Food and Agriculture Organization Joint Expert Committee on Food Additives, and Joint Meeting on Pesticide Residues) have retained the original terminology, however.

US EPA expanded the approach to include the Reference Concentration (RfC), a “safe” concentration in air analogous to the RfD, using dosimetric adjustments to the inhaled experimental animal concentration to improve the extrapolation to humans (Jarabek, 1994, 1995a, b; Jarabek et al., 1989). This yielded, for the first time, a consistent and scientifically credible replacement of part of the uncertainty factor for extrapolation from experimental animal to human, reflecting data-informed differences in biology. This transition was codified by US EPA with its publication of methods for development of inhalation RfCs (US EPA, 1994, with an update 2012g); a text on both RfDs and RfCs followed (US EPA, 2002a).

A Margin of Exposure (MOE) analysis is also often developed in chemical risk assessment for non-cancer toxicity, and, occasionally, for non-genotoxic carcinogens. A MOE is developed by dividing the NOAEL or benchmark dose (BMD) of the critical effect by the expected or measured exposures in humans. Conventionally, the default target MOE is drawn from uncertainty factors of 10 each for inter- and intra-species extrapolation, or other factors as appropriate for the critical effect of concern, to assess whether a sufficient MOE is attained to ensure safety. More recently, the MOE has

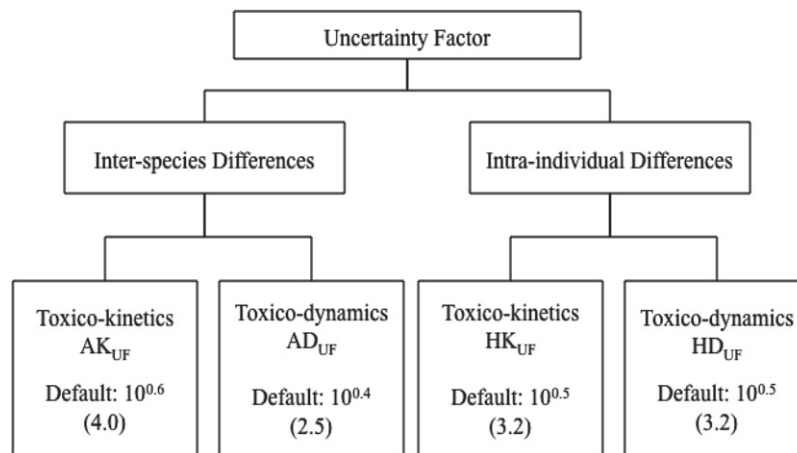
<sup>1</sup>A threshold is defined as some dose below which the probability of an individual responding is zero (Klassen, 2008—p. 23). This concept is routinely used in risk assessment. For example, recent assessments by US EPA (2012, Integrated Risk Information System, at [www.epa.gov/iris](http://www.epa.gov/iris)) include the following in the description of an RfD “The RfD is intended for use in risk assessments for health effects known or assumed to be produced through a nonlinear (presumed threshold) mode of action.”

<sup>2</sup>An adverse effect is: “a biochemical change, functional impairment, or pathologic lesion that affects the performance of the whole organism, or reduces an organism’s ability to respond to an additional environmental challenge” (US EPA, 2012e, IRIS Glossary).

<sup>3</sup>The critical effect is the first adverse effect, or its known and immediate precursor, that occurs as dose increases in the most appropriate or sensitive animal species (adapted from US EPA, 2012e).

<sup>4</sup>Reference Dose (RfD): An estimate (with uncertainty spanning perhaps an order of magnitude) of a daily oral exposure to the human population (including sensitive subgroups) that is likely to be without an appreciable risk of deleterious effects during a lifetime. It can be derived from a NOAEL, LOAEL, or benchmark dose, with uncertainty factors generally applied to reflect limitations of the data used. Generally used in US EPA’s noncancer health assessments (US EPA website accessed on 12/1/2012 at: <http://www.epa.gov/risk/glossary.htm#r>).

Figure 1. The Chemical Specific Adjustment Factor (CSAF) scheme of the International Programme on Chemical Safety (2005). The individual toxicokinetic and toxicodynamic factors are defaults to be replaced with chemical specific data, which can lead to data-derived values that are less than, equal to, or greater than the default value.



#### CSAFs

- AD<sub>UF</sub> = Uncertainty factor for animal to human differences in toxicodynamics
- AK<sub>UF</sub> = Uncertainty factor for animal to human differences in toxicokinetics
- HD<sub>UF</sub> = Uncertainty factor for human variability in toxicodynamics
- HK<sub>UF</sub> = Uncertainty factor for human variability in toxicokinetics

also been used for genotoxic carcinogens (EFSA, 2012), applying a similar approach.

Another related effort started in the early 1990s with the seminal publications of Renwick (1991, 1993). Renwick proposed replacement of the traditional 10-fold uncertainty factors addressing variability (experimental animal to human extrapolation or within human variability) with default subfactors for either toxicokinetics or toxicodynamics. In turn, these default subfactors could be replaced with chemical-specific data, when available.

As part of its harmonization<sup>5</sup> project, the WHO IPCS implemented a slightly modified Renwick approach (IPCS, 1994), followed by a decade-long series of workshops, case studies, and reviews that culminated in the development of methods for developing Chemical-Specific Adjustment Factors (CSAFs; IPCS, 2005). This work was built on numerous, often related, publications (e.g. Dourson et al., 1998; Ginsberg et al., 2002; Hattis et al., 1999; Kalberlah & Schneider, 1998; Naumann et al., 2005; Renwick, 1998a; Renwick & Lazarus, 1998b; Renwick et al., 2000, 2001; Silverman et al., 1999; Zhao et al., 1999). The IPCS effort propelled several countries to improve their process of non-cancer dose–response assessment (Health Canada by Meek et al., 1994; US EPA, 2002a, 2011e). Other groups have also

followed the IPCS paradigm, such as NSF International (Ball, 2011).

The IPCS (2005) CSAF guidance resulting from this effort specifies the approach for evaluating the adequacy of the data for replacing one or more of the four subfactors addressing variability by chemical-specific or chemical-related data. Each subfactor is independently evaluated to determine if the data are sufficient to generate a CSAF, or whether a default factor needs to be used, as shown in Figure 1.

The numerical value for a CSAF is dictated by the data and could range from less than 1 for interspecies differences to considerably more than the default subfactor for any or all of them. As a consequence, the composite uncertainty factor may be either less than or more than the usual default value, which is typically 100. If the composite factor is less than the usual default value (i.e. <100) for a particular critical effect, IPCS (2005) recommends an evaluation of other endpoints to which the usual default value might be applied, since one of these other endpoints might then become the critical effect that determines the RfD, RfC, or Tolerable Daily Intake (TDI). Although suitable data may be available only on occasion, analysis of available data on a chemical using the framework presented in the IPCS (2005) guidance provides a useful method of assessing the overall adequacy of the data for risk assessment purposes. In addition, the IPCS guidance can help direct research to identify and fill data gaps that would improve development of the safe dose. A CSAF-type approach can also be used to refine interspecies dosimetry for cancer assessments regardless of the low-dose extrapolation approach.

During this time, several other publications investigated and further developed uncertainty factors. For example, the development of a fifth area of uncertainty, that of toxicity database deficiency, was described (Baird et al., 1996; Dourson et al., 1992, 1996). US EPA (2002b) and Fenner-Crisp (2001) also published on the Food Quality Protection Act (FQPA) safety factor, showing that the hazard portion of this safety factor is addressed by proper application of this

<sup>5</sup>Harmonization as defined by International Programme on Chemical Safety (IPCS, 2005) is an understanding of the methods and practices used by various countries and organizations, acceptance of assessments that use different approaches, and a willingness to work towards convergence of these approaches or methods as a longer term goal. Achieving this goal allows comparison of information, improved understanding of the basis for exposure standards for specific chemicals in different countries (e.g. the International Toxicity Estimates for Risk (ITER) available at <http://toxnet.nlm.nih.gov/>), savings in time and expense by avoiding duplication of work, and improved science through better communication among organizations and peer review of assessments and assessment procedures. See also, for example, the Risk Information Exchange (RiskIE) available at <http://www.allianceforrisk.org/RiskIE.htm>, as a tool to facilitate collaborations and leveraging of resources.

Table 1. Continuum of effects associated with any exposure to xenobiotics reflecting a sequence of effects of differing severity (ARA, 2012).

*Adaptive effects:* This continuum starts at low dose with upstream indicators of change, or adaptive effects, where the organism's ability to withstand a challenge is enhanced. Doses associated with such effects are often referred to as No Observed Adverse Effect Levels (NOAELs). The concepts of homeostasis and hormesis are relevant here

*Compensatory effects:* As dose increases, compensatory effects occur, which enable the organism to maintain overall function without further enhancement or significant cost. Doses associated with such effects are also often NOAELs. Some of these effects might be judged to be the critical effect

*Critical effect:* As dose further increases, the critical effect is reached. This is the first adverse effect, or its known precursor, that occurs to the most [relevant or] sensitive species as the dose rate of an agent increases.<sup>a</sup> Doses associated with such effects are Lowest Observed Adverse Effect Levels (LOAELs). The highest NOAEL below this LOAEL is generally used in the dose response, and the focus is on determining this NOAEL in a sensitive population

*Adverse effects:* As dose further increases, the critical effect is exceeded, and adverse effects are manifested as biochemical changes, functional impairments, or pathologic lesions. These progressively more severe effects impair the performance of the organism, and/or reduce its ability to respond to additional challenges. At some point these adverse effects become manifestly overt and irreversible, and frank effects or clinical disease ensues

<sup>a</sup>Note that the bracketed phrase "relevant or" is important since the most relevant specie is always preferred over the most sensitive species (e.g. if data shows that the rat is more sensitive than the human, the human data are still preferred), but when such information is not available, data from the most sensitive species are chosen. Also the term "precursor" in this definition is singular, meaning the immediate precursor, not just any prior effect. This restriction is important both because it ties the concept of critical effect into common medical practice of focusing on important endpoints, and because the resulting dose response—such as an RfD—is more meaningful, since without the restriction multiple and different RfDs can be estimated.

database deficiency uncertainty factor, in conjunction with the uncertainty factor intended to address human inter-individual variability in susceptibility.<sup>6</sup> This conclusion was also reached by Dourson et al. (2002). Also, during this time Swartout et al. (1998) published an approach for developing a probabilistic description for individual and combined factors; Lewis et al. (1990) and Lewis (1993) discussed the development of adjustment factors based on data; and Pieters et al. (1998) conducted a statistical analysis of toxicity data in an evaluation of the uncertainty factor for subchronic-to-chronic extrapolation.

Recommendations that have emerged from this analysis and related efforts are:

- (1) CSAF guidelines exist for using chemical-specific or chemical-related data to characterize interspecies differences and human variability and replace default uncertainty factors. Application of these guidelines should be a standard part of developing toxicity values, as indeed they already are for many.
- (2) Scientifically based defaults are important and useful when data are insufficient to develop an adequate CSAF.
- (3) Additional factors may be utilized to account for database deficiencies such as insufficient study length (e.g. 90-day study only), absence of dose levels without adverse effects, available effects are clinically severe, or lack of data on key endpoints (e.g. developmental toxicity). Typically, these factors are applied during the derivation of a "safe dose" for data-poor chemicals.

### From critical effects to mode of action (MOA)

Risk assessment is in a state of scientific rebirth, and in order to understand the drivers behind this trend, it helps to look back over the evolution of the regulatory risk assessment process. Beginning with the 1950s, FDA and others relied on the concept of a critical target organ, or a "critical effect"

(refer to footnote 3) as described later by USEPA (e.g. Barnes & Dourson, 1988). Due in part to limitations in standard toxicity testing methods at the time, the critical effect was typically an overt toxic effect, resulting in an endpoint now referred to as an "apical effect", and often had direct clinical relevance.

As additional toxicological information was published, scientific judgment became important in distinguishing adaptive and compensatory effects from adverse effects and in identifying the critical effect and its relevant precursor effects. Table 1 shows how these effects relate to each other. Although this severity continuum has been generally accepted, a key limitation to its use is that the definition of adverse versus adaptive effects often generates controversy for individual chemicals, and often remains a challenge to toxicologists and risk assessors alike, even when an appreciation of the underlying clinical disease is considered.

A recent publication from an ILSI/HESI Committee, chartered to specifically address definitions of adverse versus adaptive, recommended the following language (Keller et al., 2012):

*Adverse Effect:* A change in morphology, physiology, growth, development, reproduction, or life span of a cell or organism, system, or (sub)population that results in an impairment of functional capacity, an impairment of the capacity to compensate for additional stress, or an increase in susceptibility to other influences;

*Adaptive Response:* In the context of toxicology, the process whereby a cell or organism responds to a xenobiotic so that the cell or organism will survive in the new environment that contains the xenobiotic without impairment of function.

This suggested language needs to be further interpreted, however, since on the face of the definition, it appears to suggest that the death of a single cell is potentially adverse, whereas redundancy within an organ would argue that it is not (Rhombert et al., 2011). A useful interpretation perhaps is to consider that an adverse effect results in the impairment of the

<sup>6</sup>Variability in exposure is not addressed by the existing uncertainty factors, but is typically addressed using conservative assumptions and high percentiles for exposure assessments.

functional capacity of the organism or higher levels of organization. This interpretation would be more consistent with the definition of adverse found in Table 1. The Key Events Dose–Response Framework (KEDRF; Julian et al., 2009) provides another means to delineate and analyze the component elements of dose–response leading up to an adverse effect and factors influencing those events (e.g. dose level and frequency, thresholds, the degree and fidelity of DNA repair, homeostatic mechanisms).

Another approach to enhance this interpretation was published by Boekelheide & Andersen (2010) because of the increasing volume of new, high-throughput data. They considered the key challenge to be the ability to distinguish between acceptable (i.e. homeostatic, adaptive and perhaps compensatory) perturbations of a pathway and excessive (i.e. critical effect) or adverse or clinically relevant effects or perturbations. Furthermore, they discussed new approaches to evaluate dose–response relationships as functions of the probabilities of biological system failure, determined in a stepwise manner through assays that measure progressive perturbation along toxicity pathways. From a biological systems perspective, it may be useful to construct such pathway analyses based on homeostasis as the organizing circuitry or network. In this manner, the dose–response of biological system failure is dictated by processes overwhelming homeostasis. From such a perspective, the “cascade of failures” of Boekelheide & Andersen (2010) ensues only when homeostasis is overwhelmed. These changes in the definition of “adverse” with the use of different types of data illustrates how one aspect of problem formulation might change the underlying biology is better understood.

That adverse effects are the product of a cascade of failures in protective processes, has also been discussed by others. Examples include error-prone or lack of DNA repair of a pro-mutagenic DNA adduct (Pottenger & Gollapudi, 2010), or failure of homeostasis and subsequent induction of fatty liver (Rhomberg, 2011). In addition, various methods have been proposed or are being developed to utilize more relevant biological data to construct models for predicting apical adverse responses, including many *in silico* approaches, molecular or mechanistic data from cells or tissues, or early biomarkers (Aldridge et al., 2006; Alon, 2007; Andersen & Krewski, 2009; Kirman et al., 2010; Yang et al., 2006). Most recently, US EPA’s ToxCast™ program has published a number of preliminary prediction models (Martin et al., 2009, 2011; Shah et al., 2011; Sipes et al., 2011).

The migration away from the conventional use of critical effects, or perhaps the integration of genomics data into the current severity scheme of Table 1, will likely require sophisticated methodologies, given the complexity of processes underlying biological pathways or networks. Prior to this, however, these newer test methods must be shown to be scientifically valid and the prediction models must be shown to have the requisite degree of scientific confidence necessary to support regulatory decisions. As discussed by Bus & Becker (2009), approaches that should be considered for method validation and predictivity include those discussed by the NRC (2007b) for toxicogenomics and the Organization for Economic Cooperation and Development (OECD) principles and guidance for the validation of quantitative structure

activity relationships (OECD, 2007). These methods and prediction models hold great promise, and significant progress continues to be made to develop and build scientific confidence in them. However, the challenges are significant. The analysis by Thomas et al. (2012a) concluded “...the current ToxCast phase I assays and chemicals have limited applicability for predicting *in vivo* chemical hazards using standard statistical classification methods. However, if viewed as a survey of potential molecular initiating events and interpreted as risk factors for toxicity, the assays may still be useful for chemical prioritization.”

A second key limitation of this severity continuum is that it focuses on apical, high-dose effects. In particular, it does not always address the problems arising from making inferences from high-dose animal toxicity studies to environmentally relevant exposures. While it is now well recognized that dose transitions and non-linearities in dose–response (Slikker et al., 2004a,b) should be integrated into extrapolation of effects from high-dose animal toxicity studies to very much lower human exposures, this was not always the case. In fact, early approaches to quantitative risk assessment, such as those described in the US EPA (1986a) cancer risk assessment guidelines, did not focus on the biology *per se*, because, at that point in time, the scarcity of mechanistic data and the limited theoretical understanding of the biological complexity of carcinogenesis made it too challenging to address these issues adequately. Although these older guidelines allowed for the use of chemical-specific data, assessments typically applied a default linear modeling approach for carcinogens when critical information about mode of action, genotoxicity or other relevant biological knowledge was unavailable, limited, or of insufficient quality. With a dearth of information, as was typical in those days of risk assessment, a general mind-set to apply defaults was pervasive. However, as described further later in this section, the growing availability of mechanistic information and increased understanding of the biology of disease processes places greater responsibility on risk assessors to utilize all the available effects data (from homeostatic, adaptive, compensatory, critical, adverse and clinical outcomes) within the focus and limitations identified in the problem formulation. Unfortunately, in some US government programs the default approaches have been so ingrained that it has proven very difficult to incorporate this newer, biologically based information and methods.

Although the US EPA (1986a) cancer risk assessment guidelines and related early US EPA publications for non-cancer toxicity (Barnes & Dourson, 1988) emphasized defaults, they provided a framework for considering integration of data obtained from different study types. Thus, these guidelines were intended to be sufficiently flexible to accommodate new knowledge and assessment methodologies as such methods were developed. One advantage of these first steps was to reduce the required effort in hazard identification by concentrating on a single, manageable piece of information: the critical effect. By focusing the risk assessment on a single critical effect and setting risk values to be protective for that critical effect, it was presumed that exposed populations would be protected against all other apical effects of concern, as such effects would require higher doses to manifest. The US EPA (1986a) guidelines also allowed for the incorporation

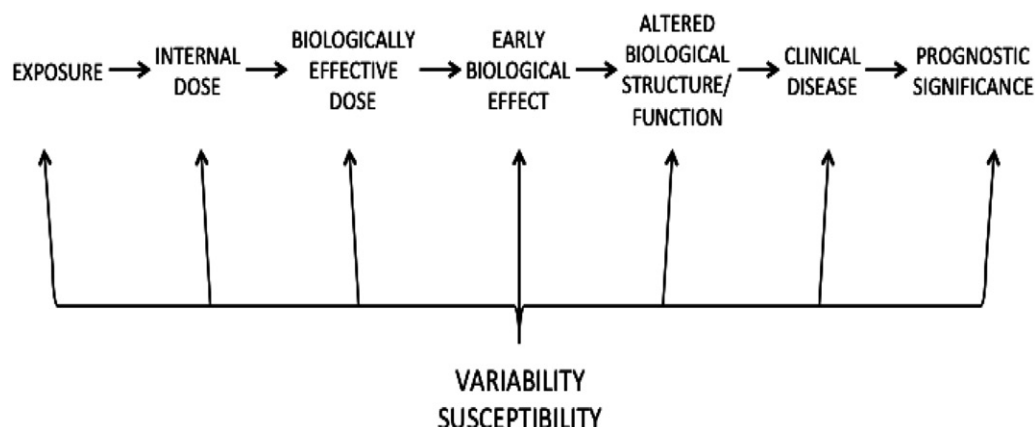


Figure 2. Series of steps that occurs between exposure and the effect of clinical disease and prognostic significance. Adapted from Schulte (1989).

of mechanistic data in place of default extrapolation procedures despite the fact that such data were rarely available at the time.

Schulte (1989) and NRC (1989) opened a new chapter in risk assessment by providing a structure for considering the series of steps that occurs between exposure and the toxic effect (Figure 2) [adapted from Schulte, 1989]. These steps delineate areas for acquisition of data illuminating how a chemical might cause the observed effects. Specific and quantifiable biomarkers related to each specific step can be used to replace the “black box” between exposure and effect. The NRC (1989) report classified biomarkers as markers of exposure, markers of effect, and markers of susceptibility.

Schulte’s pathologic progression diagram laid the foundation in part for work by US EPA, IPCS, and others attempting to determine the type and level of information needed to use non-default approaches. A key concept in this evolution was a focus on MOA rather than mechanism of action. While a mechanism of action reflects the detailed, molecular understanding of a biological pathway, the MOA characterizes a more general understanding of how the chemical acts. The MOA is defined as a sequential series of key events, with a key event being defined as an empirically observable and quantifiable precursor step that is a necessary (but not necessarily sufficient) element of the MOA or is a biologically based marker for such an element. Determination of dose–response for key events is an important aspect of establishing an MOA. The US EPA cancer guidelines (USEPA, 1996, 2005) are key documents describing the potential applications of MOA data. Specifically, these guidance documents recommend using data as the starting point where possible (data before defaults), and focusing upon assessment of weight of evidence, with the goal of applying the MOA approach to all appropriate data.

During the same time period, a number of projects at ILSI and IPCS further developed the MOA approach, initially for carcinogens (Sonich-Mullin et al., 2001), and then for non-carcinogens (Seed et al., 2005), with particular emphasis on using MOA information to evaluate HR, culminating in the development of the mode of action/human relevance framework (MOA/HRF) (Meek et al., 2003; IPCS, 2006; Sonich-Mullin et al., 2001). In this framework (Figure 3) [from WHO IPCS, 2007], one first uses the modified Hill criteria to

determine whether the data are sufficient to determine the acting MOA in experimental animals. If the MOA is established in an experimental animal model, the HR framework goes on to evaluate whether the HR of the MOA can be excluded, first based on fundamental, qualitative differences in key events between animals and humans, and then based on quantitative differences.

Both qualitative and quantitative differences in MOA and resulting responses should be considered. If the HR cannot be excluded, then the MOA is assumed to be applicable to humans, and then quantitative toxicokinetic or toxicodynamic data can be used to replace defaults with CSAFs. Qualitatively, if a MOA is determined to not be relevant to humans, then that MOA can be excluded from the human health risk assessment (e.g. male rat kidney tumors caused by alpha 2u-globulin nephropathy—Hard et al., 1993). Other MOAs or endpoints caused by that chemical of concern can then be evaluated to determine whether they are relevant to humans. One clear strength of this approach is that both chemical-specific information and a general understanding of biology and physiology are used to address fundamental questions regarding the MOA, dose–response, and toxicity of a specific chemical. In the future, advanced mechanistic-based molecular screening approaches may increasingly reveal quantitative differences between human-based assays and animal-based assays that may improve the accuracy of risk assessments.

The MOA/HRF continues to be refined as experience is gained in its application. For example, it is now recognized that absolute responses to the framework questions are not needed. Instead, the MOA/HRF questions provide a structure for describing the degree of confidence and uncertainties associated with application of available data in risk assessments (Meek & Klaunig, 2010). Another new element of this approach is recognition of the importance of “modulating factors,” such as polymorphisms, pre-existing disease states, and concurrent chemical exposures, which can affect susceptibility to risk (Meek, 2008). Detailed examples of modulating factors provided by Meek (2008) included differences in the presence and activity of enzymes in biotransformation pathways, competing pathways of biotransformation, and cell proliferation induced by coexisting pathology. The MOA/HRF can also be used to aid in identifying populations or life stages that may have increased susceptibility.

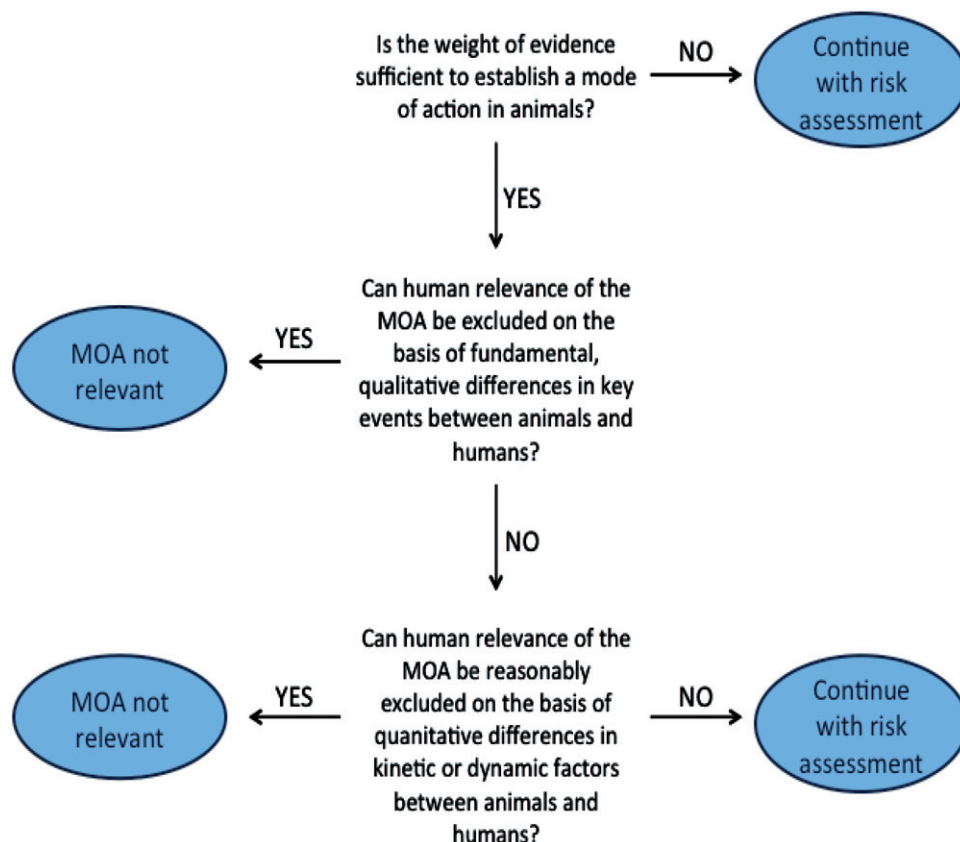


Figure 3. The mode of action/human relevance framework (MOA/HRF). Adapted from WHO (2007).

Recently, the KEDRF was developed as an extension of the MOA/HRF (Boobis et al., 2009; Julien et al., 2009). This framework considers the dose–response and variability associated with each key event to better understand and potentially quantitate the impact of each of these factors on the risk assessment as a whole. For example, in considering mutation as a potential key event, one considers whether mutation is likely an early rate- or dose-limiting step, or whether it is secondary to other effects, such as cytotoxicity and compensatory cell proliferation (Meek & Klaunig, 2010). Furthermore, the KEDRF can be used to compare the dose necessary to elicit the key event(s) in relation to doses actually experienced in real-world exposures.

A number of advantages exist to the use of MOA data and the MOA/HRF/KEDRF or a similar framework. First, in-depth assessments can be conducted with it. Second, consideration of MOA issues can aid in developing and refining research strategies (Meek, 2008). For example, as an example of the interplay between problem formulation and biological considerations, discussions between risk assessors and research scientists can improve the efficiency of risk assessments by focusing resources on tiered and/or targeted approaches that are more efficient and reduce animal use (Meek, 2008; Meek & Klaunig, 2010), as envisioned by NRC (2007a). Focusing on earlier, potentially more sensitive biological endpoints that represent key events will facilitate the use of data directly from environmentally relevant human exposures, and/or the use of *in vitro* model systems using human-derived tissues or cells. Such approaches would not only have increased relevance to human physiology, they also

would have the potential to be used in high- or medium-throughput formats. Carmichael et al. (2011) noted that even today, standard test protocols do not always provide the information needed to support a MOA analysis. Better incorporation of MOA information is facilitated by the increased understanding of the multiple ways in which such data can be incorporated into risk assessment, as well as in the early focus on hazard characterization.

Another advantage to the use of MOA data is that extensive research over the last 30 years can be reviewed to test the default linear and non-linear low dose extrapolation procedures. This has been done and non-linear MOAs for chemical carcinogens appear to be more scientifically justified, when compared with the default linear procedure, in a number of instances (Boobis et al., 2009; Cohen & Arnold, 2011). Cohen & Arnold (2011) conclude that for non-DNA reactive carcinogens, “[i]n each of these instances studied in detail, the carcinogenic effect is because of an increase in cell proliferation. This can either be by a direct mitogenic effect (involving hormones and/or growth factors) or can be because of toxicity and regeneration.” They further state that knowledge garnered from research on mode of action that distinguishes DNA-reactive from non-DNA-reactive carcinogens “... forms the basis for the distinction of potential risks to humans in regulatory decision making.”

The KEDRF provides one structure for describing the degree of confidence and uncertainties associated with reliance on such knowledge and data in *lieu* of the linear default. For many substances that produce cancer in laboratory animal studies, even for those that may cause point

mutations in genotoxicity assays, assessors are failing to objectively describe the evidence for alternatives to linear low-dose extrapolation (Boobis et al., 2009; Cohen & Arnold, 2011; Swenberg et al., 2011). Determining the most appropriate model(s) and approach(es) for regulatory risk assessment for a specific substance will be guided by statutes, policies and scientific knowledge. For either DNA-reactive or non-DNA-reactive substances, the statistical characterization of the low-dose dose–response relationship for tumorigenesis *in vivo* would require prohibitively large numbers of lab animals. Therefore, our expanding knowledge of the pathogenesis of cancer (Cohen & Arnold, 2011; Hanahan & Weinberg, 2000) and experimental data sets which evaluate MOAs in the carcinogenic process (e.g. biomarkers of DNA damage, cell-proliferation, pathway addiction, clonal expansion, DNA-methylation, tumor suppressor gene expression) are key to profiling substances according to patterns of biological responses. These profiles can be compared to the profiles of prototypical chemical carcinogens, and in this manner, empirical dose–response data of a substance can be integrated with knowledge of MOA, both broadly and for the specific chemical, to enhance the scientific basis of risk assessment. Scientific knowledge of MOA today is simply too advanced to support undue reliance upon a default-driven system for evaluating carcinogenic or non-carcinogenic risks to humans.

Although the use of MOA has been growing substantially, scientific hurdles to increased regulatory acceptance of MOA-based approaches remain. Such hurdles include lack of empirical data to define the shape of the dose–response curve at low, environmentally relevant exposures, and incomplete knowledge of what constitutes “scientific sufficiency” for the purpose of defining a MOA and its presumed low-dose dose–response relationship for regulatory risk assessment purposes. A 2009 workshop to address this general issue of integration of MOA into risk assessment made the following recommendations (Carmichael et al., 2011):

- Establish a group of experts from a variety of backgrounds to generate a database of accepted MOAs and to identify minimum data requirements needed to characterize a chemical’s MOA;
- Generate guidance documents describing the appropriate means by which MOA data can be incorporated into chemical risk assessments;
- Promote a shift in current risk assessment practices to focus on hazard characterization using MOA data; also, identify what information could be provided by standard toxicity tests to inform the MOA evaluation;
- Utilize a tiered and flexible framework to collect and apply MOA data to assessments;
- Develop predictive methods for MOA based on evaluation of early key events;
- Optimize use of data collected in human trials or clinical studies; and
- Globally harmonize MOA terminology.

Furthermore, the NRC report entitled *Toxicity Testing in the 21st Century: A Vision and a Strategy* (NRC, 2007a), aims to harness MOA information ultimately to generate a battery of *in vitro* tests to evaluate chemical-specific toxicity, concomitantly reducing the need for whole animal studies

and focusing research on biological pathways. A number of other consensus reports and guidelines also support measures to increase the focus on MOA as the central organizing principle, and use of *in vitro* data to reduce animal use, although the general consensus of these reports is that animal testing would not be eliminated, at least not in the near term (Carmichael et al., 2011; NRC, 2009; US EPA, 2005; WHO IPCS, 2007).

For the most part, with the exception of genotoxicity assays, the application of *in vitro* data directly into risk assessment is in its infancy. For such data to be effectively incorporated into hazard characterization and dose–response assessment, they will have to be vetted against traditional approaches and harmonized with clinical practice. As such approaches are proven valid over time, they are expected to streamline the risk assessment process itself, allowing for more efficient assessments and read-across interpretations among chemical groups that share MOAs. In addition, the use of cell culture models to address risk assessment will ultimately reduce the need for studies conducted in animals, minimizing animal usage to more focused, MOA studies. Moreover, such approaches facilitate prioritization of chemicals based on anticipated risk to human health.

Recommendations that have emerged from this analysis and related efforts are:

- (1) Focus must shift away from identification of only a toxicant-induced apical effect (critical effect) towards identification of a sequence of key events/MOA as the organizing principle for risk assessment.
- (2) Development and acceptance of standardized definitions are essential for adverse effect, adaptive response, and MOA, and for how such data may be integrated with clinical knowledge in order to improved risk assessment.
- (3) Identification of early, driving key events in toxicity/biological pathways will be necessary to apply MOA as the organizing principle. To effectively analyze such key events, a refined context of the dose necessary to elicit them is needed in relation to doses actually experienced from real-world exposures.

## Low-dose extrapolation: transition from defaults to mode of action (MOA) understanding

### Underlying assumptions

As noted above, the default approach for non-cancer dose–response assessment assumes a threshold for an adverse effect and uses uncertainty factors to estimate a safe dose, while current default dose–response approaches for cancer assessment often assumes that no threshold exists, resulting in a linear extrapolation from the observed animal data to low doses, especially if genotoxicity has been demonstrated or not adequately ruled out. Although recent publications have demonstrated many examples of *in vitro* and *in vivo* non-linear or even threshold dose–response for gene mutations and micronucleus formation induced by DNA reactive chemicals (Bryce et al., 2010; Doak et al., 2007; Gocke & Müller, 2009; Gollapudi et al., 2013; Pottenger et al., 2009), the linear dose–response approach has traditionally been selected based on this assumption of no threshold, and the resulting linear extrapolation is considered the most

conservative. These two divergent approaches for dose–response assessment reflect not only these different assumptions, but also the fundamental nature of the cellular damage and the body’s ability to handle such damage. Different regulatory policies follow.

Linear extrapolation for cancer, for example, is based on a stochastic assumption: that the potential for critical damage to DNA is a matter of chance, and that this probability depends only on dose in a linear relationship, so that a doubling of dose results in a directly proportional increase in the chance of critical DNA damage (Dourson & Haber, 2010; US EPA, 1976; US EPA, 1986a; US EPA, 2005). It further assumes that a single heritable change to DNA can induce malignant transformation, leading to cancer. Other factors, such as an individual’s repair capacity or a chemical’s toxicokinetics are assumed to be independent of dose, so that the risk per unit dose is constant in the low-dose range. As further discussed by Dourson & Haber (2010), low-dose linear extrapolation is a convenient health-protective approach. However, factors such as the efficiency of DNA repair, rate of cell proliferation, and chemical-specific toxicokinetics indicate that even if the dose–response for cancer is linear at low (environmentally relevant or lower) doses, the slope of that line is likely to be lower than the slope of the line extrapolating from the animal tumor data to zero (Swenberg et al., 1987). Cohen & Arnold (2011) note that DNA-reactive carcinogens produce “strikingly non-linear dose–response” curves, due in part to an acceleration of damage, or lack of repair at higher doses when compared to lower doses. Fortunately, the new biological tools available now and in the near future will be capable of experimentally testing the assumption that DNA-reactive substances demonstrate linearity at low doses. For example, recent work on directly DNA-reactive radiation effects demonstrate non-linear dose–response for a variety of molecular events such as base lesions, micronuclei, homologous recombination, and gene expression changes following low-dose exposures (Olipitz et al., 2012). Outcomes of these and other experiments challenge the need for maintaining the dichotomy between cancer and non-cancer toxicities, and between genotoxic and non-genotoxic chemicals with respect to potential carcinogenic risk to humans at environmentally relevant exposures.

In contrast to mutagenic effects initiated by chemicals directly interacting with DNA, the safe dose assessment for non-cancer endpoints<sup>7</sup> assumes that cells have many molecules of each protein and other targets. And, thus, damage to a single molecule is not expected to lead to a damaged cell. In fact, if damage to one molecule of a single cell were sufficient to cause it to die, redundancy in the target organ would mean that the cell’s death is not adverse, as more fully explicated by Rhomberg et al. (2011). Based on the redundancy of target molecules and target cells, together with the capacity for repair, regeneration or replacement, these adverse effects are assumed to have a threshold. In addition, the sigmoidal dose–response curve often produced

by quantal data (apical adverse effects) in linear space occurs as a result of the variability in individual responses and underlying genomic plasticity, reflecting differences in sensitivity to a given chemical. In the highly unlikely occurrence of no differences in sensitivity among individuals, the population dose–response would be expected to be a step function, with no response below a certain dose, and up to 100% response above that dose. Such responses are seldom, if ever, seen, thus supporting the assumption that the sigmoidal response curve for quantal data is influenced by individual variability in response.

### Using mode of action (MOA) information

At the molecular level, log dose–response curves are typically sigmoidal because the response is the result of the ligand binding (reversibly) to a single receptor site and thus directly proportional to receptor binding (law of mass action; Balakrishnan, 1991). Furthermore, when response is mediated by a cascade of messengers following the initial binding of the ligand to the receptor, as long as the subsequent responses are the result of the messenger molecule binding to a single binding site, according to the law of mass action, the dose–response curve will be the same sigmoid shape as the initial receptor binding dose–response.

However, depending on the mechanism, the shape of the dose–response curve for the ultimate toxic effect (the apical effect) will vary, and as Conolly and Lutz (2004) note, “Actions of a toxic agent in an organism are multifaceted, the reaction of the organism accordingly is pleiotropic, the dose–response is the result of a superimposition of all interactions that pertain.” Thus, it is important to articulate the MOA and analyze the corresponding key events. This may be particularly true in carcinogenesis, where, “six essential alterations in cell physiology collectively dictate malignant growth: self-sufficiency in growth signals, insensitivity to growth-inhibitory (antigrowth) signals, evasion of programmed cell death (apoptosis), limitless replicative potential, sustained angiogenesis and tissue invasion and metastasis” (Hanahan & Weinberg, 2000).

While these different default approaches reflect different underlying assumptions, there is general agreement on the preference for use of MOA to inform the dose–response assessment. Both of the recent NRC reports (NRC, 2007a, 2009) acknowledge the importance of using MOA to inform risk assessment, including improving animal to human extrapolations (or removing the need for such extrapolation) and characterizing the impact of human variability on these extrapolations. In fact, many recent guidance documents and committee recommendations point to the importance of incorporating MOA data into risk assessment approaches (e.g. Seed et al., 2005; US EPA, 2005; WHO IPCS, 2007).

To the degree that differences exist among these recommendations, they occur mostly in the application of MOA data in risk assessment. According to NRC (2007a), US EPA (2005) and others, MOA is the central driver, upon which decisions about dose–response assessment should be based. In contrast, the NRC (2009), while stating that MOA evaluations are central, recommends the use of low-dose linear

<sup>7</sup>As noted elsewhere in this text, the same assumption applies to cancer resulting from MOAs other than interaction with DNA.

extrapolation as a default for non-cancer toxicity, and as the preferred default approach for harmonizing<sup>8</sup> cancer and non-cancer dose–response assessment. Both of these NRC (2009) recommendations appear to run counter to toxicological and biological principles (Rhombert et al., 2011). Furthermore, these recommendations fail to address differences in the assumptions underlying the two default extrapolation procedures as discussed above.

Perhaps not surprisingly, these recommendations of NRC (2009) also run counter to other recommendations to establish harmonized default risk assessment paradigms (Crump et al., 1997, 1998; IPCS, 2006; Meek, 2008; NRC, 2007a; US EPA, 2005). In fact, based on these publications, and given what is now known about toxicity mechanisms, DNA damage and repair, and homeostasis, a biological case can be made that the preferred default approach would be to harmonize non-cancer and cancer assessments using the KEDRF approach, or if insufficient information exists for the KEDRF, then on the basis of expected thresholds or non-linearities for adverse effect.

For example, Rhombert et al. (2011) published a critique of the NRC (2009) report emphasizing that low-dose linearity for non-cancer effects was the exception, not the rule, and therefore, not an adequate basis for a universal default position. These authors counter the NRC (2009) recommendation that low-dose linear is the scientifically justified default based on (1) considerations of distributions of inter-individual variability, (2) interaction with background disease processes, and (3) undefined chemical background additivity. Rhombert et al. (2011) show: (1) that the “additivity-to-background” rationale for linearity only holds if it is related to a specific MOA, which has certain properties that would not be expected for most non-cancer effects (e.g. there is a background incidence of the disease in the unexposed population that occurs via the same pathological process as the effects induced by exposure); (2) that variations in sensitivity in a population tend to only broaden, not linearize, the dose–response relationship; (3) that epidemiological evidence of purported linear or no-threshold effects at low exposures in humans, despite non-linear exposure–response in the experimental dose range in animal testing for similar endpoints, is most likely attributable to exposure

measurement error rather than a true linear association. In fact, only implausible distributions of inter-individual variation in parameters governing individual sigmoidal response could ever result in a low dose linear dose–response. The last NRC (2009) justification (i.e. undefined chemical background additivity) is also discounted as a justification by Dourson & Haber (2010), since such background is better addressed by standard risk assessment methods for chemical mixtures. Indeed the dual NRC (2009) recommendations to use low-dose linear extrapolation as a default for non-cancer toxicity, and as the preferred default approach for harmonization, work against US EPA’s mixtures guidelines that recommend adding individual chemical dose–response assessments together in the form of a HI.

Of the two different NRC (2007a, 2009) approaches to harmonization of cancer and non-cancer risk assessment, the approach recommended by the NRC (2007a) and others, to harmonize using MOA as the organizing principle, appears scientifically stronger. By relying on MOA as the harmonizing principle, the focus is more on the relevant biology rather than mathematical or statistical tools. A useful example of this preferred approach to cancer and non-cancer risk assessments based on US EPA (2005) guidance is found in the published propylene oxide (PO) cancer MOA risk assessment (Sweeney et al., 2009). PO is a nasal respiratory irritant, and the PO cancer MOA is a complex series of biological responses driven by PO induction of severe, sustained GSH depletion in target rat nasal mucosa, which leads to nasal respiratory epithelial cell proliferation concomitant with significant irritation, and eventually to nasal tumors. The induction of cell proliferation and nasal irritation is identified as the critical key event and has been characterized as having a “practical threshold”; thus the harmonized cancer/non-cancer risk assessment relies on determination of exposure limits low enough to protect against induction of nasal irritation, which will then protect from both non-cancer and cancer effects (Sweeney et al., 2009). In this case, the MOA based on sustained cell proliferation was used to inform the risk assessment despite the fact that PO is capable of causing genetic damage. The authors concluded that the MOA data were sufficient in this case to justify a threshold model for dose–response assessment, instead of the default linear no-threshold model. Several authoritative bodies have cited this article and have accepted the threshold MOA for PO-induced cancer, including the European Union Scientific Committee on Occupational Exposure Levels (SCOEL, 2010) and the German MAK Commission (MAK, 2012).

### *Dose–response modeling*

Linear extrapolation is a default policy choice that is intended to be health-protective in the face of uncertainties. Its use in this regard is considered to protect public health. However, a number of demonstrated non-linearities or thresholds exist in the biology of cancer, even for chemicals acting *via* a mutagenic MOA. Such non-linearities or thresholds can occur as a result of numerous biological processes, including uptake, transport, metabolism, excretion, receptor binding and DNA repair and other cellular defense mechanisms. Thus,

<sup>8</sup>Harmonization of cancer and noncancer endpoints is clearly not a novel concept, given the impetus of former committees and organizations. However, the NRC (2009) specifically recommends that harmonization should be focused around dose–response and proposes three conceptual models described as (CM1): nonlinear individual response, low-dose linear population response with background dependence (i.e. overall linear, non-threshold response from which a slope factor is most appropriate); (CM2): low-dose nonlinear individual and nonlinear population response, low-dose response independent of background (i.e. a threshold response for which a reference dose is most appropriate); and (CM3): low-dose linear individual and linear population dose–response (i.e. a linear, non-threshold response from which a slope factor is most appropriate). The report further clarifies that low-dose linear refers to the slope in the low-dose region, and “it does not mean that the dose–response relationship is linear throughout the dose range between zero dose and high doses.” The approach has been described as “piecewise linear,” to capture the idea of different slopes in different regions. The NRC (2009), however, does not provide further guidance on how to characterize the low-dose slope as something other than the linear slope between a point of departure in the experimental dose range and the origin.

when considering the entire dose–response curve, linear extrapolation from the apical endpoint of cancer needs to be carefully considered in relationship to the available evidence regarding the MOA and the resulting shape of the dose–response curve (Dourson & Haber, 2010; Hattis, 1990; Slikker et al., 2004a). The emphasis on MOA, then, is not determining whether non-linearities or thresholds exist, but more on how best to capture modern knowledge and understanding of the underlying biology related to the chemical's dose–response curve and its ultimate relevance to adverse health outcomes.

Slikker et al. (2004a, 2004b) described this issue with the idea of dose-dependent transitions. Not unlike the NAS (2009), they noted that quantal dose–response curves can often be thought of as “serial linear relationships,” due to the transitions between mechanistically linked, saturable, rate-limiting steps leading from exposure to the apical toxic effect. To capture this biology, Slikker et al. (2004a) recommended that MOA information could be used to identify a “transition dose” to be used as a point of departure for risk assessments instead of a NOAEL/LOAEL/BMDL. This transition dose, if suitably adjusted to reflect species differences and within human variability, might serve as a basis for subsequent risk management actions.

The key events dose–response framework (KEDRF; Boobis et al., 2009; Julien et al., 2009) further incorporates a biological understanding by using MOA data and information on shape of the dose–response for key events to inform an understanding of the shape of the dose–response for the apical effect. This applies both to fitting the dose–response curve to the experimental data in the range of observation as well as for extrapolation. Advantages of the KEDRF approach include the focus on biology and MOA, consideration of outcomes at individual and population levels, and reduction of reliance on default assumptions. The KEDRF focuses on improving the basis for choosing between linear and non-linear extrapolation, if needed, and, perhaps more importantly, extending available dose–response data on biological transitions for early key events in the pathway to the apical effect; in short, another way to extend the relevant dose–response curve to lower doses.

Biologically based modeling can be used to yet further improve the description of a chemical's dose–response. PBPK modeling predicts internal measures of dose (a dose metric), which can then be used in a dose–response assessment of a chemical's toxicity, and so can directly capture the impact of kinetic non-linearities on tissue dose. This information can be used for such applications as improving interspecies extrapolations, characterization of human variability, and extrapolations across exposure scenarios (Bois et al., 2010; Lipscomb et al., 2012). PBPK models can also be used to test the plausibility of different dose metrics, and thus the credibility of hypothesized MOAs. Recent guidance documents and reviews (IPCS, 2010; McLanahan et al., 2012; USEPA, 2006c) provide guidance on best practices for characterizing, evaluating, and applying PBPK models. Additional extrapolation to environmentally relevant doses can be addressed with PBPK modeling.

Biologically based dose–response (BBDR) modeling adds a mathematical description of the toxicodynamic effects of

the chemical to a PBPK model, thus linking predicted internal/tissue dose to toxicity response. Perhaps the best-known BBDR model is that for nasal tumors from inhalation exposure to formaldehyde (Conolly et al., 2003), which builds from the Moolgavkar-Venzon-Knudson (MVK) model of multistage carcinogenesis (Moolgavkar & Knudson, 1981). The formaldehyde BBDR predicts a threshold, or at most a very shallow dose–response curve, for the tumor response despite evidence of formaldehyde-induced genetic damage. MVK modeling of naphthalene, focusing on tumor type and joint operation of both genotoxic and cytotoxic MOAs, is illustrative of an MOA approach that can be taken to quantitatively evaluate risk (Bogen, 2008). Further, Bogen (2008) demonstrates how to quantify the potential upper-bound, low dose, non-threshold (genotoxic) contribution to increase in tumor risk. Such approaches may be useful for quantitative risk evaluations for a number of substances where two or more MOAs may be involved. Such approaches are encouraged by guidelines for cancer risk assessment (EPA, 2005).

Other approaches that are less data-intense can use chemical-specific or chemical-related information to extend the dose–response curve into the range (or near the range) of the exposures of interest. These approaches allow one to use mechanistic data more directly to evaluate dose–response, without having to evoke default approaches of linear or non-linear extrapolation. Such biologically informed empirical dose–response modeling approaches have the goal of improving the quantitative description of the biological processes determining the shape of the dose–response curve for chemicals for which it is not feasible to invest the resources to develop and verify a BBDR. An advantage of these approaches is using quantitative data on early events (biomarkers) to extend the overall dose–response curve to lower doses using biology, rather than being limited to the default choices of linear extrapolation or uncertainty factors.

In one demonstration of this sort of approach, Allen et al. (submitted), outlined a hypothesized series of key events describing the MOA for lung tumors resulting from exposure to titanium dioxide (TiO<sub>2</sub>), building on the MOA evaluation of Dankovic et al. (2007). Allen et al. used a series of linked “cause-effect” functions, fit using a likelihood approach, to describe the relationships between successive key events and the ultimate tumor response. This approach was used to evaluate a hypothesized pathway for biomarker progression from a biomarker of exposure (lung burden), through several intermediate potential biomarkers of effect, to the clinical effect of interest (lung tumor production). Similar work has been published by Shuey et al. (1995) and Lau et al. (2000) on the developmental toxicity of 5-fluorouracil.

Another approach to biologically informed empirical dose–response modeling was demonstrated by Hack et al. (2010), who used a Bayesian network model to integrate diverse types of data and conduct a biomarker-based exposure-dose–response assessment for benzene-induced acute myeloid leukemia (AML). The network approach was used to evaluate and compare individual biomarkers and quantitatively link the biomarkers along the exposure-disease continuum. This work provides a quantitative approach for linking changes in biomarkers of effect both to exposure

information and to changes in disease response. Such linkage can provide a scientifically valid point of departure that incorporates precursor dose–response information without being dependent on the difficult issue of a definition of adversity for precursors.

Even less computationally intensive mechanistic approaches are possible. For example, Strawson et al. (2003) evaluated the implications of exceeding the RfD for nitrate, for which the critical effect is methemoglobinemia in infants. They based their analysis on information on the amount of hemoglobin in an infant's body and the amount of nitrate required to oxidize hemoglobin to an adverse level; extrapolation was not needed, since data are available for the target population (human infants) in the adverse effect range.

Physiologically based pharmacokinetic/pharmacodynamic (PBPK/PD) models are also useful for improving the biological description of the shape of the dose–response into low-dose regions without the full complexity of a BBDR. Price et al. (2011) and Hinderliter et al. (2011) combined PBPK/PD modeling and Monte Carlo modeling to develop a source-to-outcome model that provides a quantitative description of the relationship between the amount of dietary residues of the insecticide chlorpyrifos in food, and the impact of the exposures on inhibition of brain and red blood cell cholinesterase in exposed populations, including consideration of sensitive populations. The model identified a dose that does not cause a biologically meaningful change in cholinesterase inhibition (a critical precursor key event), and thus a dose where adverse effects are not expected, even for individuals who are at increased susceptibility due to other stressors. While cholinesterase inhibition may not have a biological threshold, this is a good example of a threshold for an adverse effect, since these small measureable changes in cholinesterase, a key event metric, are not toxicologically or clinically meaningful. So while this example shows that clinical parameters may differ from the mean or reflect perturbations of physiological homeostasis, they are not “adverse effects.”

Other approaches use statistical methods to better characterize the dose–response curve. For example, categorical regression has been used to estimate the risk for non-cancer endpoints (Teuschler et al., 1999), and the signal-to-noise crossover dose (SCND) has been suggested as an alternative point of departure for extrapolation (Sand et al., 2011; see also comment by Chiu et al., 2012). SNCD indicates where the increased risk is greater than the background variability, and is defined as “the dose where the point estimate of additional risk is equal to or, alternatively,  $0.67 \times$  the (absolute) difference between the upper and lower bound of a two-sided 90% CI on absolute risk at that dose” (Sand et al., 2011).

#### *Areas of future growth in understanding MOA*

The explosive growth of systems biology, fueled to a significant degree by the NRC (2007a) report, offers great promise for improving the characterization of dose–response curves, and aiding in the movement away from defaults. The biologically informed empirically based approaches provide a useful bridge to incorporating mechanistic information, but

the ideal envisioned by NRC (2007a) is to be able to base risk assessments primarily on *in vitro* studies in human cells. This approach would allow one to test environmentally relevant doses in the appropriate test species, and use targeted approaches to consider human sensitivity, coupled with targeted *in vivo* testing for resolving specific issues.

Many challenges remain prior to achievement of the NRC (2007a) vision, including phenotypic anchoring of early biomarker changes (Thomas et al., 2012b; Waters & Fostel, 2004), as well as the development of methods for evaluating effects on organ systems, rather than at the tissue level. However, progress has been made in developing tools for using early biomarkers and ‘omics measures. For example, pathway-based toxicogenomic analyses have found substantial correlation between transcriptional benchmark dose (BMD) values from subchronic studies and those for apical non-cancer and cancer endpoints (Thomas et al., 2012b).

At the heart of the NRC (2007a, 2009) and other reports are many recommendations on the use of data instead of defaults to guide the risk assessment process, a goal upon which all can agree. The challenge remains to incorporate an understanding of the MOA into the broad range of fit-for-purpose applications to risk assessment, rather than reliance on default procedures, and many approaches described above, such as the CSAF, MOA/HR and KEDRS frameworks, are instructive here.

Recommendations that have emerged from this analysis and related efforts are:

- (1) Harmonization of cancer and non-cancer dose–response assessments should be conducted on the basis of MOA understanding, using such frameworks as the MOA/HR and KEDRS.
- (2) Systems biology approaches will be useful in better characterizing the biology of low, environmentally relevant dose–responses and their relevance to clinical findings.
- (3) Additional work is needed on dose–response methods and models that better capture the biology across the full range of the dose–response, particularly in the low dose region.

#### **Cumulative risk and mixtures**

A well-recognized problem in human health risk assessment has been that while the estimation of risk from single contaminants is fairly well established, human exposures are nearly always to chemical mixtures, or to multiple chemicals in a sequential fashion. In fact, we are exposed to many thousands of chemicals daily, most of which are natural rather than synthetic. In addition to exogenous exposures, certain substances are formed endogenously, and the specific mix of chemical exposures varies from day to day depending on our environment and activity. Moreover, the increased sensitivity of modern analytical methods allows us to measure simultaneously more chemicals at lower concentrations in human fluids and tissues than ever before. Thus, detection of numerous substances in biomonitoring a single individual is not unexpected, as simultaneous exposure to chemicals in our environment is the rule, not the exception. Numerous early

attempts have been made to deal with these issues, but both the methodology for evaluating potential risks from such mixtures and indeed even the mixtures risk assessment nomenclature are varied and can be stultifying.

Guidelines for mixtures risk assessment have been developed by a number of authoritative organizations (e.g. ACGIH, 2011; ATSDR, 2001a, 2011b; Meek et al., 2011; US EPA, 1986b, 2000b). The first, and most straight forward, but highly limited approach, is to directly assess the dose–response for the mixture of concern (US EPA, 1986b, 2000). A second, related approach is to directly assess the dose–response for a sufficiently similar mixture (US EPA, 2000b). A third approach involves the dose–response assessment of individual chemicals within the mixture, and combining the assessments of individual chemicals based on either independent action or dose addition, depending on what is known about the MOAs for the various chemicals in the mixture. These assessments can be modified to make appropriate adjustments for multiple and differing chemical interactions, including consideration of similar and dissimilar kinetics and dynamics. It is with this latter approach that the NRC's (2009) recommendation for harmonization of cancer and non-cancer approaches is in direct opposition. NRC (2009) states that undefined background additivity caused by co-exposure to similarly acting chemicals or coexisting disease processes support implementation of its recommended default linear approach. However, the US EPA's third approach to chemical mixtures depends upon evaluation of *individual* chemical dose–response assessments. If an individual chemical dose–response assessment were to incorporate background exposures to other similarly acting chemicals, as suggested by NRC, then the resulting HI would reflect double counting or would not be needed (Dourson & Haber, 2010).

Evaluating interactions among chemicals with differing concentrations within mixtures can be challenging. The three approaches described above have the most utility for product safety or environmental assessments. The preferred approach is the one in which the final product or contamination that will reach humans is tested directly (i.e. the first approach mentioned above). This approach provides clarity on the conditions under which the product may be used safely or the contamination is without risk. Historically, however, it is the third approach described above, where individual chemical risks are evaluated and compared, that received the most use in regulatory decisions. In these cases, independent action has been generally assumed for substances believed to cause toxicity through dissimilar modes of action (ATSDR, 2001a, b; USEPA, 2000). Under the independent action assumption, so long as exposure to each component of a mixture occurs at its safe dose or below, no toxicological effects of the mixture would be expected. There is a substantial body of scientific literature to support independent action at low exposure levels (e.g. Borgert et al., 2012; US EPA, 2000b). Although exposure levels will typically differ by orders of magnitude, in both workplace settings for industrial hygiene practices (ACGIH, 2011) and in screening assessments for evaluating potential exposures from hazardous waste sites (US EPA, 1989), dose addition among individual chemicals has been likewise used in the absence of information on specific mixtures. A common approach to dose addition is the HI, which is the

sum of the Hazard Quotient (HQ) for each chemical in the mixture ( $\text{Hazard Quotient}_i = \text{Exposure}_i \div \text{Safe Dose}_i$ ). This approach assumes that exposure to a mixture of substances, in which each component is at a subthreshold dose for toxicity, could result in an adverse health effect, when the summed exposures (weighted by their safe doses) exceed an HI of 1, as estimated in an iterative fashion.

It can be seen that adding the HQs for each chemical to develop a HI is very useful for screening purposes, because it eliminates from consideration situations that are considered to be without any risk. However, when a HI of 1 is exceeded with such an approach, an adverse effect should not be presumed. In such cases, one could either manage the presumed risk at that point, or further refine either the exposure or hazard assessment. The US EPA directs analysts to refine the HI approach by segregation of chemicals by similar toxic effect or similar MOA, a common US EPA practice at Superfund sites. These multiple HIs are then each compared to a value of 1, and if none are exceeded then the situation is considered to be without any risk. Such an iterative approach is key in applying the various approaches described above. Another technique is the Toxicity Equivalency Factor (TEF) where the potencies of a set of similarly related chemicals are assumed related to each other, or related to a sentinel chemical.

The HI approach has generally been recognized as a conservative, health-protective application, both because the estimations of maximum exposures and safe doses each in turn employ conservative assumptions, and because scientific justification for dose additivity is robust only in cases where chemicals induce the similar toxic effect by the same MOA and exposure doses are either close to or in the operative range of the dose–response for the individual chemicals in the mixture. Thus, application of the HI approach to simultaneous exposures to multiple chemicals for which the chemicals do not induce the similar toxic effect or do not act by the same MOA will overestimate potential risk. It is for this reason that the HI approach is most appropriately applied as a screening level method.

Following the passage of the FQPA in 1996, which required US EPA to (1) determine the cumulative effects of pesticides that have a common mechanism of toxicity and (2) ensure that there is a reasonable certainty that no harm from aggregate exposure to the pesticides, US EPA devoted considerable resources to develop and apply specific procedures to conduct aggregate and cumulative risk assessment (<http://www.epa.gov/oppfead1/trac/science/>).<sup>9</sup> For cumulative risk for pesticide exposures, the US EPA's framework employs dose addition when the chemicals (usually grouped with other structurally related chemicals) cause the same effect *via* a common mechanism/mode of action. To date, US EPA's Office of Pesticide Programs has conducted cumulative risk assessments for five groups of pesticides:

<sup>9</sup>In this context, aggregate risk refers to exposure to the same chemical from multiple routes, while cumulative exposure refers to exposure to multiple chemicals, multiple routes. Due to the substantial inconsistency in how these and other terms are used, Meek et al. (2011) recommended that the “aggregate” and “cumulative” terms be replaced by more explicit terms such as “single chemical, multiple routes” and “multiple chemicals, multiple routes,” respectively.

organophosphates, n-methyl carbamates, triazines, chloroacetanilides and pyrethrins/pyrethroids (<http://www.epa.gov/oppsrrd1/cumulative/>).

A probabilistic method to evaluate multiple simultaneous exposures to chemicals acting by similar and dissimilar modes of action has been developed (NRC, 2004). This approach, in which dose addition is used for substances with a common mechanism and independent action is used for substances with different modes of action, clearly shows how important it is to base a cumulative risk assessment upon knowledge of mode of action.

In contrast perhaps, the NRC (2008b) report “Phthalates and Cumulative Risk Assessment: The Task Ahead” recommended applying dose addition to all chemicals that produce “common adverse outcomes.” However, without a clear definition of common adverse outcome, this recommendation might suggest that the initial screening level HI approach is preferred, with little emphasis on the iterative nature of subsequent approaches or a clearer understanding of underlying MOA. Borgert et al. (2012) recently showed that the underlying assumptions and analysis in support of this “common adverse outcomes” recommendation of the NRC (2008b) are useful only primarily as a coarse screening level assessment, and that refined approaches are needed once one considers larger numbers of chemicals. Moreover, Borgert et al. (2012) point out that one should consider the relative

exposures between the laboratory animal NOAEL and the estimated human exposure when analyzing the independent action of the toxicants. This can be seen in one sense as supporting the current iterative approach.

Based in part on these prior deliberations, a unifying integrating framework, presented in Figure 4, has been published for evaluating the risk of combined exposure to multiple chemicals (Meek et al., 2011). Based on a workshop of the WHO/IPCS, the framework specifies a four-tiered iterative approach that integrates hazard and exposure assessments for risk-based decision making. In the IPCS framework, if the screening level evaluation based on the assumption of dose addition for all chemicals is adequate, that is if the HI is equal to or less than a value of 1 or if the margin between the overall exposure and an appropriate hazard marker is considered sufficient, no further action would be required. However, if the HI or margin of exposure raises concern, the next step can be generation of additional data, refinement of the exposure and/or hazard assessment (where the latter would include MOA at Tier 2), or a risk management decision. The WHO/IPCS tiered approach has the advantage of not only building on previous guidelines, but also incorporating new thinking on Toxicity Testing in the 21st Century (NRC, 2007a) in that such testing is likely to expand our understanding and use of MOA information as recommended by NRC (2009).

### Problem Formulation for Combined Exposure Assessment

- *What is the nature of the exposure?*
- *Is exposure likely, taking into account the context?*
- *Is there a likelihood of co-exposure within a relevant timeframe?*
- *What is the rationale for considering compounds in an assessment group?*



### Example Tiered Exposure and Hazard Considerations: Mixture or Component Based

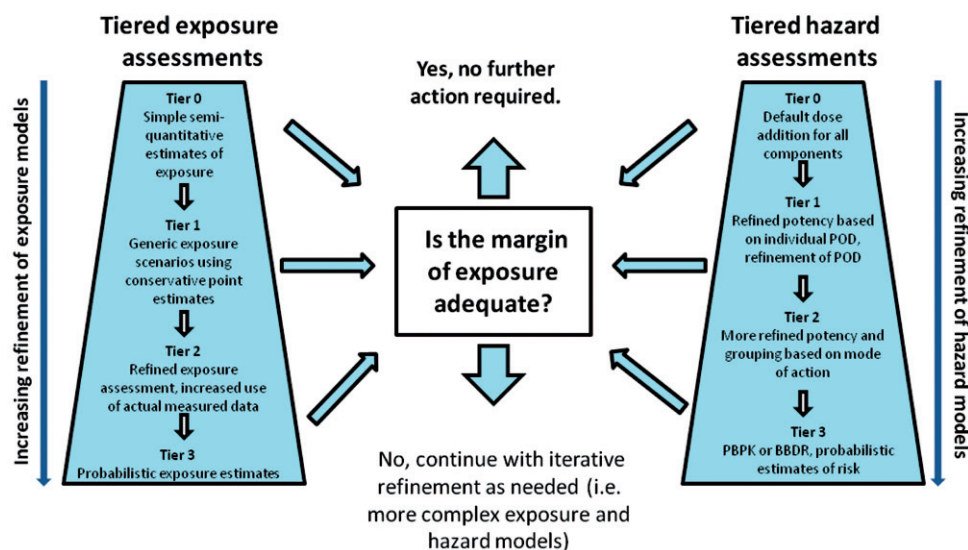


Figure 4. Unifying integrating framework for evaluating the risk of combined exposure to multiple chemicals. From Meek et al. (2011) (Reprinted from Regulatory Toxicology and Pharmacology, Volume 60 (2011) S1–S14; by Bette Meek, Alan R. Boobis, Kevin M. Crofton, Gerhard Heinemeyer, Marcel Van Raaij, and Carolyn Vickers, entitled Risk assessment of combined exposure to multiple chemicals: A WHO/IPCS framework, with permission from Elsevier.).

Other authors have also considered adaptations to this WHO/IPCS framework. For example, Price & Han (2011) show how Maximum Cumulative Ratio (MCR), the ratio of the cumulative toxicity received by an individual from exposure to multiple chemicals to the largest toxicity from a single chemical, can be used as part of the WHO/IPCS Tier 1 and Tier 2 assessments. The MCR approach of Price & Han (2011) predicts that, for the vast majority of mixture exposures, the key determinant of toxicity resides in the single most toxic agent in the mixture.

Recommendations that have emerged from this analysis and related efforts are

- (1) Approaches to the risk assessment of chemical mixtures should be iterative.
- (2) A HI summation method based on all adverse outcomes offers a simplistic approach that will adequately protect public health against adverse effects. However, this approach is not applicable beyond screening.
- (3) The tiered framework of IPCS (Meek et al., 2011) integrates relevant and scientifically appropriate prior information and should be used as a template for future work. This iterative approach guides refinement of the exposure assessment and/or use of common MOA to replace the screening HI approach.
- (4) Different problem formulations allow different uses of the iterative IPCS framework.

## Biomonitoring

Biomonitoring programs provide an opportunity to better associate real-world exposures (internal doses) to the dose–response and MOA data used in a risk assessment. This is accomplished by comparing an internal equivalent to the safe dose (or other dose response value) to the levels detected in biomonitoring studies. Advanced analytical methods in human biomonitoring can now provide accurate identification and quantification of dozens of substances in reasonable sample volumes at the individual level; and human biomonitoring programs at the national, state, and international level have advanced similarly. Prominent population-based programs include:

- CDC's National Biomonitoring Program (<http://www.cdc.gov/biomonitoring/>);
- California Biomonitoring Program (<http://www.cdph.ca.gov/programs/Biomonitoring/Pages/default.aspx>);
- Minnesota's Biomonitoring Pilot Program (<http://www.health.state.mn.us/divs/hpcd/tracking/biomonitoring/index.html>);
- Canadian Health Measures Survey Biomonitoring Program (<http://www.hc-sc.gc.ca/ewh-semt/contaminants/human-humaine/glance-resume-eng.php>);
- Consortium to Perform Human Biomonitoring on a European Scale (<http://www.eu-hbm.info/cophes>); and
- German Human Biomonitoring Program (<http://www.umweltbundesmat.de/gesundheit-e/monitor/index.htm>).

However, as was clearly pointed out by the National Research Council (NRC, 2006), “In spite of its [human biomonitoring] potential, tremendous challenges surround the use of biomonitoring, and our ability to generate biomonitoring data has exceeded our ability to interpret what the data

mean to public health.” The NRC panel, recognizing that methods for interpreting human biomonitoring data in a health risk context and for communicating this interpretation were equally important, recommended developing tools to enhance both the scientific interpretation of biomonitoring data and to address these communication challenges (NRC, 2006).

To meet this need for methods, both forward dosimetry (Hays et al., 2007) and reverse dosimetry (Liao et al., 2007) approaches were developed that use pharmacokinetic tools to convert an applied dose for humans into an internal concentration (and *vice versa*). By applying pharmacokinetic tools to existing chemical risk assessments, external safe doses, such as reference doses or tolerable daily intakes, or other dose response values, can be converted to corresponding biomarker concentrations in blood or urine, and with such a conversion, a comparison can then be made to the actual concentration measured by a human biomonitoring study.

As a result of this initial work, an expert panel developed guidance for driving and documenting “Biomonitoring Equivalents” or “BEs,” for use in interpreting human biomonitoring results in a health risk context (Hays et al., 2008). The BE is defined as the concentration of a chemical or metabolite in a biological medium (blood, urine, human milk, etc.) that is consistent with an existing exposure guidance value such as a tolerable daily intake (TDI) or reference dose (RfD). Subsequently, the BE approach has been extensively refined and expanded (reviewed in Becker et al., 2012; Hays & Aylward, 2009), and BE values have been derived for approximately 80 chemicals, including persistent organic compounds (e.g. dioxins, hexachlorobenzene, and DDT), approximately 40 volatile organic compounds, phthalates and phenols (bisphenol A, triclosan and several phthalates), certain pyrethroid pesticides, and for selected brominated flame retardant compounds (Angerer et al., 2011). Scientists from the private sector, CDC/ATSDR and US EPA have recently collaborated to compare BEs to measured biomarker concentrations for approximately 130 substances in the CDC's National Exposure Report to provide a risk assessment perspective on the significance of the levels of these substances found in the US population (Aylward et al., 2013).

However as mentioned above, without methods to interpret biomonitoring results in a risk context, risk assessors and risk managers (or, the general public, for that matter) cannot distinguish the significance of the exposures. In light of these significant advances in developing tools for interpreting human biomonitoring data and the recognition and guidance from authoritative organizations such as the Centers for Disease Control and Prevention that the mere detection of a substance does not equate to illness or injury, a communication strategy has been developed for BEs by LaKind et al. (2008a). Key communication issues from these authors include:

- Developing a definition of the BE that accurately captures the BE concept in lay terms;
- Communicating comparisons between population biomonitoring data and BEs;
- Communicating to individuals and groups the significance of biomonitoring data that exceed BEs for a specific chemical;

- Describing the level of confidence in chemical-specific BEs; and
- Developing key requirements for effective communication with health care professionals.

While the risk communication literature specific to biomonitoring is sparse, many of the concepts developed for traditional risk assessments apply, including transparency and discussions of confidence and uncertainty. Best communication practices dictate use of the most credible scientific analysis, which for human biomonitoring translates into interpreting and communicating results in a responsible manner using tools such as BEs. With BEs, the measured biomonitoring data can be quantitatively interpreted within the context of a KEDREF/MOA evaluation. Interpreting biomonitoring in a risk context maximizes its value and impact by empowering health professionals to communicate results to individuals and groups in terms of their health concerns. BEs also enable risk managers and the public to decide if and when additional management actions are warranted, and permit risk-based approaches for prioritizing resources. Interpretations based only on consideration of presence are still being published (e.g. Woodruff et al., 2011), but while full disclosure of information is to be commended, doing so without a corresponding communication strategy that informs the public on relevance should be actively discouraged.

As with any human study, biomonitoring studies need to comply with the Common Rule (DHHS, 1991), which requires informed consent, minimization of avoidable risks, and independent ethical review by an Institutional Review Board (IRB). This review includes the complete study protocol, consent forms and communications materials. One of the challenges in biomonitoring studies pertains to dissemination of results to study participants, particularly when existing knowledge is limited as to the potential health significance of the levels of specific substances detected in an individual's specimen. As Harrison (2008) has pointed out, the bioethical "...principle of autonomy supports the 'right to know,' but the principles of beneficence, non-maleficence and veracity seem to support nondisclosure." Foster & Agzarian (2007) suggest reporting results to individuals for substances for which "there is credible evidence linking exposure with adverse health effects in the human population" but not for those substances for which "human health risks and intervention levels are unknown." The development of BEs has expanded the basis for interpreting human biomonitoring results in a health risk context. Clearly, approaches for communicating results to individuals should be an integral part of the protocol of a well-designed biomonitoring study.

As biomonitoring techniques have advanced, so too have research studies aimed at better understanding how potential health outcomes relate to environmental exposures. In particular, the availability of the NHANES data sets, which include metrics of health status and biomonitoring levels, has stimulated numerous cross sectional analyses exploring potential associations between exposures and health. However, detection of such associations is far from establishing causality since such studies are unable to ascertain the temporal sequence of exposure and outcome (LaKind et al., 2008b, 2012).

Tangentially related to BEs, are cross sectional epidemiology studies that are often reported in mainstream media as "evidence" of effects in humans. While most researchers are careful to state that such studies cannot establish cause and effect, they often to not report effects of multiple comparisons. Moreover, such limitations are often overlooked by the media. Therefore, even for cross sectional studies, application of the Hill criteria should be considered both by investigators when interpreting their studies, and by peers when reviewing studies for publication in scientific journals. These criteria include, among others, strength and consistency of associations, temporality of exposure and effect, specificity, biological plausibility, and dose-response.

Finally, in many cross sectional and case control studies it seems as if a chemical's potential MOA is not evaluated. This can be rectified by use of knowledge from human clinical findings and toxicity studies in laboratory animals. For example, Zhao et al. (2005) used knowledge of clinical findings and dose-response data in laboratory animals to determine the likely MOA for chlorpyrifos. They then compared this MOA to human epidemiological results of Whyatt et al. (2004) to show that it was not biologically feasible to conclude that the levels of chlorpyrifos in a study of newborns in New York city were causally related to low birth weights, as was asserted. Other epidemiology findings confirmed the analysis by Zhao et al. (2005). A general journal practice of encouraging the publication of the underlying data supporting key conclusions of studies can help support independent analyses that investigate apparent contradictions between studies in experimental animals and epidemiology investigations (see, e.g. Souza et al., 2007; Vines et al., 2013). Such practice will also aid in the interpretation of BEs.

Recommendations that have emerged from this analysis and related efforts are:

- (1) Analytical methods in human biomonitoring now provide accurate quantification of many substances in biological samples; biomonitoring programs exist at the national, state, and international levels and provide a unique and valuable snapshot of population exposures to chemicals in our environment.
- (2) Biomonitoring equivalents and supporting methods for interpreting human biomonitoring data in a health risk context now exist and should be used. Case studies published in the open literature are available for further guidance.
- (3) Interpreting human biomonitoring data in a public health risk context vastly increases the value of population-based biomonitoring programs by allowing risk managers to easily compare population risks from chemical exposures across a broad range of compounds.
- (4) Epidemiological studies investigating potential associations of biomonitoring results with health status or health outcomes should include the development of communication materials in their protocols and subject to IRB review.
- (5) Publications of cross sectional and case control studies should explicitly include a discussion of the effects of multiple comparisons; analysis of consistency of

associations, temporality, specificity, biological plausibility, and dose–response; and an evaluation of a chemical's potential MOA.

## Discussion

Multinational groups of scientists have labored long and hard to develop risk assessment frameworks that incorporate the best science, allow the use of more data in order to better reflect the relevant biology and clinical importance, and promote harmonization of risk assessment approaches across a broad range of toxicological responses. Through debate and discussion, a general consensus is emerging from these efforts.

First, the concept of problem formulation, and its necessary planning and scoping as a prelude to risk assessment development, is generally embraced by all organizations that evaluate health impacts of chemicals. Different risk management decisions can be, and are being, based on different problem formulations. A risk management decision requiring setting priorities for testing among a large number of substances appropriately dictates a different risk assessment approach when compared with decisions for setting clean-up levels in soil at waste sites proposed for residential redevelopment. Importantly, while risk management input on problem formulation is essential in order for risk assessment scientists to develop useful information, this upfront identification of risk management options should not be seen as changing, subverting, corrupting, or circumventing the scientific process.

Second, CSAF guidelines exist for using chemical-specific or chemical-related data to characterize interspecies differences and human variability and replace default uncertainty factors. Although scientifically based defaults are important and useful when data are insufficient to develop an adequate CSAF, the consideration of these factors should be a standard part of developing toxicity values in dose response assessment.

Third, scientific data, in particular those that inform the identification of MOAs, are increasingly providing a central organizing principle for any assessment. US EPA and IPCS guidelines on topics such as MOA/HRF, and KEDRF exist to aid assessors in integrating MOA information into risk assessments for both cancer and non-cancer health endpoints. Such data are also now being routinely integrated into the development of safe doses, and CSAF guidelines specifically exist to do this for non-cancer, and appropriate cancer, health endpoints. However, scientifically based defaults are important and useful when data on MOA and/or CSAFs are either absent or insufficient to support risk assessment decisions.

Fourth, harmonization of cancer and non-cancer dose–response assessments is now increasingly being accomplished on the basis of MOA understanding, and relevant biology and clinical significance, using guidelines described above (e.g. US EPA, 2012f for chloroform and Dourson et al., 2008 for acrylamide). Although existing default procedures remain different between cancer and non-cancer dose–response based on current scientific understanding of stochastic processes

(for cancer) and individual variability (for non-cancer), a need might exist for yet another harmonized default procedure for these two types of toxicity as suggested by NRC (2009), but if so, these two underlying assumptions will first need to be harmonized. In fact, as knowledge of biological pathogenesis, homeostasis, and dose-dependent transitions continue to be understood and integrated, it seems that a low-dose, non-linear, biological threshold,<sup>10</sup> dose–response approach might emerge as the choice for this default harmonization, if it is needed at all.

Ultimately, for safety determinations, the goal should be to use data over defaults by integrating knowledge of biological systems with data on chemical interactions with such systems (kinetics and dynamics), to characterize dose–responses. The MOA/HRF and KEDRF provide a means through which data and lines of evidence from human epidemiological studies, animal toxicity studies, and mechanistic investigations can all be integrated to determine potential risks to humans at environmentally relevant levels of exposure. These frameworks should become the standard operating procedure for all risk assessments, as indeed they already are for many.

Fifth, approaches to the risk assessment of combined exposures, such as chemical mixtures, are iterative. The tiered framework of IPCS (Meek et al., 2011) builds upon prior guidelines of ACGIH, US EPA, and others, and integrates relevant and scientifically appropriate prior information. It should be used as a template for current and future risk assessment work. Efforts to flesh out higher hazard tiers of the IPCS framework in terms of MOA understanding have been published (Borgert et al., 2012). Keeping exposures below the acceptable HI, or a similarly defined construct, using evaluations that are developed using increasingly informed tiered approaches, will adequately protect public health against adverse effects.

Sixth, analytical methods in human biomonitoring now provide accurate quantification of many substances in biological samples. The fact that biomonitoring programs at the national, state, and international levels are currently collecting such data necessitates the availability and application of methods for interpreting human biomonitoring data in a health risk context, which now exist with associated case studies for further guidance. These methods can be based on an up-to-date understanding of MOA and dose–response, which can aid in the thoughtful and appropriate communication of the health implications of biomonitoring data. Such communication is particularly important for when such data are collected and shared on an individual or group or location-specific basis.

In an effort to further clarify and unify these and the many other on-going discussions in the risk assessment community, the Alliance for Risk Assessment (ARA, 2013) has organized a series of multi-collaborator meetings, entitled “Beyond Science and Decisions: From Problem Formulation to Dose–Response Assessment.” This continuing series of meetings is led by an expert panel, which guides discussions on the ever-evolving use of biological

<sup>10</sup>Biological threshold is defined here as a biologically meaningful, either adverse or clinical, increase over background.

data in dose–response assessment *via* case studies by building from the methods framework presented in NRC (2009). The panel has developed a practical risk methods framework, which charts a path forward for the risk assessment community for differing problem formulations, providing illustrative case studies on alternative methodologies for dose–response assessments for each of these different problem formulations. The work of this ARA group has been summarized by Meek et al. (in press).

In summary, we all aspire to improved chemical risk assessment leading to better-informed risk management to promote the protection of human health and the environment within a framework of sustainable development. Within this aspiration, many expert groups have weighed in on improvements to dose–response assessment, just one aspect of the still standard, 4-step risk assessment paradigm of the NRC (1983, 2009). Many of the stated recommendations are mutually supportive, but others are conflicting. In this review article, we strive to give a concise synthesis that would allow readers to chart a course through these differing committee recommendations for advancing human dose–response assessment.

While this review covers a wide survey of advisory committee recommendations with the goal of improving chemical risk assessment policies and practices, it is not a top to bottom evaluation of the current state of risk assessment in the US or globally. Nor does this review analyze the progress by US EPA with their new initiatives such as NexGen (<http://www.epa.gov/risk/nexgen/>) and ToxCast (<http://www.epa.gov/nct/toxcast/>) or specifically address in detail other chemical hazard and risk assessment programs such as those within FDA, NIEHS/NTP, CPSC, etc. Nevertheless, virtually all of the recommendations discussed above are applicable to the design and conduct of such programs. Importantly, we have found that the series of multi-collaborator meetings organized by the ARA, which brings together experts and practitioners from the public, private and not-for-profit sectors, is an efficient and effective means to broaden and deepen scientific discourse on key topics for advancing risk assessment. Often, the venues created by agencies to engage stakeholders are too tightly controlled, too limited in time, and too narrowly focused by the agencies themselves to obtain meaningful substantive input. Expansion of the ARA model beyond dose–response, to encompass other aspects of risk assessment, holds great promise for fostering harmonization of approaches and improving the policies and practices of risk assessment throughout state and federal programs in the US and globally.

## Acknowledgements

We thank the authors of the many committee reports whose work we briefly summarize and analyze. We also appreciate the many insights resulting from the Alliance for Risk Assessment workshop series (ARA, 2013), which are based in part on the thoughtful consideration of the appropriate evolution of risk assessment by authors of Toxicity Testing in the 21st Century (NRC, 2007a) and Science and Decisions (NRC, 2009). Finally, the authors sincerely thank Annie Jarabek and Julie Fitzpatrick of US EPA for their many helpful discussions and insights.

## Declaration of interest

The affiliation of the authors is as stated on the cover page. Toxicology Excellence for Risk Assessment (TERA) is a non-profit organization with a mission to support the protection of public health by developing, reviewing and communicating risk assessment values and analyses; improving risk methods through research; and, educating risk assessors, managers, and the public on risk assessment issues. The American Chemistry Council is organized to promote innovation, safety, policies, products and technologies associated with its member chemical industries. The Dow Chemical Company specializes in developing chemicals to improve agriculture, lifestyle, energy use, and infrastructure and transportation. The Texas Commission on Environmental Quality strives to protect its state's public health and natural resources consistent with sustainable economic development. Penelope Fenner-Crisp, a retired senior EPA official, consults on risk assessment issues for a variety of organizations, including governments and NGOs.

The authors thank the American Chemistry Council and the organizations of the individual authors for support to develop this text. Each of these organizations had an opportunity to review this text as part of their internal clearance. However, this article is exclusively the work product of the authors and does not necessarily represent views or policies of the authors' employers or sponsors.

## References

- ACGIH (American Conference of Governmental and Industrial Hygienists). (2011). TLVs and BEIs based on the documentation of the "Threshold limit values for chemical substances and physical agents and biological exposure indices". p. 78.
- Aldridge BB, Burke JM, Lauffenburger DA, Sorger PK. (2006). Physicochemical modeling of cell signaling pathways. *Nat Cell Biol*, 8, 1195–203.
- Allen B, Maier A, Willis A, Haber LT. (2013). Use of early effect biomarker data to enhance dose–response models of lung tumors in rats exposed to titanium dioxide. Submitted.
- Alon U. (2007). Network motifs: theory and experimental approaches. *Nat Rev Genet*, 8, 450–61.
- ARA (Alliance for Risk Assessment). (2013). Beyond science and decisions: from problem formulation to dose response. Available from: [http://www.allianceforrisk.org/ARA\\_Dose-Response.htm](http://www.allianceforrisk.org/ARA_Dose-Response.htm) [last accessed 10 March 2013].
- Andersen ME, Krewski D. (2009). Toxicity testing in the 21st century: bringing the vision to life. *Toxicol Sci*, 107, 324–30.
- Angerer J, Aylward LL, Hays SM, et al. (2011). Human biomonitoring assessment values: approaches and data requirements. *Int J Hyg Environ Health*, 241, 348–60.
- ATSDR. (2001a). Guidance manual for the assessment of joint toxic action of chemical mixtures. Atlanta, GA: Agency for Toxic Substances and Disease Registry, Public Health Service, U.S. Department of Health and Human Services.
- ATSDR. (2001b). Guidance manual for the preparation of an interaction profile. Atlanta, GA: Agency for Toxic Substances and Disease Registry, Public Health Service, U.S. Department of Health and Human Services.
- Aylward LL, Kirman CR, Schoeny R, et al. (2013). Evaluation of biomonitoring data from the CDC National exposure report in a risk assessment context: perspectives across chemicals. *Environ Health Perspect*, 121, 287–94. doi: 10.1289/ehp.1205740.
- Bachmann J. (2007). Will the circle be unbroken: a history of the U.S. national ambient air quality standards. *J Air Waste Manage Assoc*, 57, 652–97.
- Baird JS, Cohen JT, Graham JD, et al. (1996). Noncancer risk assessment: probabilistic characterization of population threshold doses. *J Hum Ecol Risk Assess*, 2, 79–102.

- Balakrishnan N, ed. (1991). Handbook of the logistic distribution. New York, NY: CRC Press.
- Ball G. (2011). Personal communication with M. Dourson, Toxicology Excellence for Risk Assessment (TERA), Cincinnati, Ohio, August 23.
- Barnes DG, Dourson ML. (1988). Reference dose (RfD): description and use in health risk assessments. *Regul Toxicol Pharmacol*, 8, 471–86.
- Becker RA, Hays SM, Robison S, Aylward LL. (2012). Development of screening tools for the interpretation of chemical biomonitoring data. *J Toxicol Article ID 941082*, 10 pages. Available from: <http://www.hindawi.com/journals/jt/2012/941082/> [last accessed 10 March 2013].
- Birnbaum L, Damstra T, Hart J, et al. (2001). Integrated risk assessment: a report prepared for the WHO/UNEP/ILO International Programme on Chemical Safety. WHO/IPCS/IRA/01/12.
- Boekelheide K, Andersen ME. (2010). A mechanistic redefinition of adverse effects – a key step in the toxicity testing paradigm shift. *Altox*, 27, 243–52.
- Bogen KT. (2008). An adjustment factor for mode-of-action uncertainty with dual-mode carcinogens: the case of naphthalene-induced nasal tumors in rats. *Risk Anal*, 28, 1033–51.
- Bois FY, Jamei M, Clewell HJ. (2010). PBPK modeling of inter-individual variability in the pharmacokinetics of environmental chemicals. *Toxicology*, 278, 256–67.
- Boobis AR, Daston G, Preston J, Olin S. (2009). Application of key events analysis to chemical carcinogens and noncarcinogens. *Crit Rev Food Sci Nutr*, 49, 690–707.
- Borgert CJ, Sargent EV, Casella G, et al. (2012). The human relevant potency threshold: reducing uncertainty by human calibration of cumulative risk assessments. *Regul Toxicol Pharmacol*, 62, 313–28.
- Bryce SM, Avlasevich SL, Bemis JC, et al. (2010). Miniaturized flow cytometric in vitro micronucleus assay represents an efficient tool for comprehensively characterizing genotoxicity dose-response relationships. *Mutat Res*, 703, 191–9.
- Bus JS, Becker RA. (2009). Toxicity testing in the 21st century: a view from the chemical industry. *Toxicol Sci*, 112, 297–302.
- Carmichael N, Bausen M, Boobis AR, et al. (2011). Using mode of action information to improve regulatory decision-making: an ECETOC/ILSI RF/HESI workshop overview. *Crit Rev Toxicol*, 41, 175–86.
- Chiu WA, Guyton KZ, Hogan K, Jinot J. (2012). Approaches to human health risk assessment based on the signal-to-noise crossover dose. Correspondence. *Environ Health Perspect*, 120, a264.
- Clegg DJ. (1978). Toxicology basis of the ADI — present and future considerations. In: Frehse H, Geissbuhler H, eds. *Pesticide reviews*. Oxford, England: Pergamon Press, 74–7.
- Cohen SM, Arnold LL. (2011). Chemical carcinogenesis. *Toxicol Sci*, 120, S76–92.
- Committee on Biological Effects of Ionizing Radiation Committee (BEIR). (2006). Health risks from exposure to low levels of ionizing radiation: BEIR VII Phase 2. Nuclear and Radiation Studies Board, Division on Earth and Life Studies, National Research Council of the National Academies. Washington, DC: The National Academies Press.
- Conolly RB, Kimbell JS, Janszen D, et al. (2003). Biologically motivated computational modeling of formaldehyde carcinogenicity in the F344 rat. *Toxicol Sci*, 75, 432–47.
- Conolly RB, Lutz WK. (2004). Nonmonotonic dose-response relationships: mechanistic basis, kinetic modeling, and implications for risk assessment. *Toxicol Sci*, 77, 151–7.
- Crump KS, Clewell III HJ., Andersen ME. (1997). Cancer and non-cancer risk assessments should be harmonized. *Human Ecol Risk Assess*, 3, 495–9.
- Crump KS, Clewell III HJ., Andersen ME. (1998). Cancer and non-cancer risk assessments should be harmonized. *Comment Toxicol*, 6, 277–82.
- Dankovic D, Kuempel E, Wheeler M. (2007). An approach to risk assessment for TiO<sub>2</sub>. *Inhal Toxicol*, 19, 205–12.
- Daston G, Faustman E, Ginsberg G, et al. (2004). A framework for assessing risks to children from exposure to environmental agents. *Environ Health Persp*, 112, 238–56.
- DHHS (Department of Health and Human Services). (1991). Common Rule. National Institutes of Health, Office for Protection from Research Risks Title 45 Code of Federal Regulations Part 46. Protection of Human Subjects.
- Doak SH, Jenkins GJ, Johnson GE, et al. (2007). Mechanistic influences for mutation induction curves after exposure to DNA-reactive carcinogens. *Cancer Res*, 67, 3904–11.
- Dourson ML, Charnley G, Scheuplein R. (2002). Differential sensitivity of children and adults to chemical toxicity: II. Risk and regulation. *Regul Toxicol Pharmacol*, 35, 448–67. Available from: <http://www.tera.org/Publications/Dourson%202002.pdf> [last accessed 10 March 2013].
- Dourson ML, DeRosa CT. (1991). The use of uncertainty factors in establishing safe levels of exposure. In: Krewski D, Franklin C, eds. *Statistics in toxicology*. New York, NY: Gordon and Breach Science Publishers, 613–27.
- Dourson ML, Felter SP, Robinson D. (1996). Evolution of science-based uncertainty factors in noncancer risk assessment. *Regul Toxicol Pharmacol*, 24, 108–20.
- Dourson ML, Haber LT. (2010). Linear low-dose extrapolations. In: Hsu CH, Stedeford T, eds. *Cancer risk assessment, chemical carcinogenesis, hazard evaluation, and risk quantification*. Hoboken, NJ: John Wiley & Sons, Inc., pp. 615–35.
- Dourson ML, Hertzberg R, Allen B, et al. (2008). Evidence-based dose response assessment for thyroid tumorigenesis from acrylamide. *Regul Toxicol Pharmacol*, 52, 264–89.
- Dourson ML, Knauf LA, Swartout JC. (1992). On reference dose (RfD) and its underlying toxicity data base. *Toxicol Ind Health*, 8, 171–89.
- Dourson ML, Maier A, Meek B, et al. (1998). Re-evaluation of toxicokinetics for data-derived uncertainty factors. *Biol Trace Element Res*, 66, 453–63. Available from: <http://www.tera.org/Publications/Boron1998.pdf> [last accessed 10 March 2013].
- Dourson ML, Stara JF. (1983). Regulatory history and experimental support of uncertainty (safety) factors. *Regul Toxicol Pharmacol*, 3, 224–38.
- EFSA (European Food Safety Authority). (2012). Statement on the applicability of the Margin of Exposure approach for the safety assessment of impurities which are both genotoxic and carcinogenic in substances added to food/feed. *EFSA J*, 10, 2578.
- Fenner-Crisp P. (2001). The FQPA 10x safety factor—how much is science? how much is sociology? *HERA*, 7, 107–16.
- Foster WG, Agzarian J. (2007). Reporting results of biomonitoring studies. *Anal Bioanal Chem*, 387, 137–40.
- Ginsberg G, Hattis D, Sonawane B, et al. (2002). Evaluation of child/adult pharmacokinetic differences from a database derived from the therapeutic drug literature. *Toxicol Sci*, 66, 185–200.
- Gocke E, Müller L. (2009). In vivo studies in the mouse to define a threshold for the genotoxicity of EMS and ENU. *Mutat Res*, 678, 95–100.
- Gollapudi BB, Johnson GE, Hernandez LG, et al. (2013). Quantitative approaches for assessing dose-response relationships in genetic toxicology studies. *Environ Molec Mutagenesis*, 54, 8–18.
- Hack CE, Haber LT, Maier A, et al. (2010). A Bayesian network model for biomarker-based dose response. *Risk Anal*, 30, 1037–51.
- Hanahan D, Weinberg RA. (2000). The hallmarks of cancer. *Cell*, 100, 57–70.
- Hard GC, Rodgers IS, Baetcke KP, et al. (1993). Hazard evaluation of chemicals that cause accumulation of alpha 2u-globulin, hyaline droplet nephropathy, and tubule neoplasia in the kidneys of male rats. *Environ Health Perspect*, 99, 313–49.
- Harrison M. (2008). Applying bioethical principles to human biomonitoring. *Environ Health*, 7, S8.
- Hattis D. (1990). Pharmacokinetic principles for dose rate extrapolation of carcinogenic risk from genetically active agents. *Risk Anal*, 10, 303–16.
- Hattis D, Banati P, Goble R, Burmaster DE. (1999). Human interindividual variability in parameters related to health risks. *Risk Anal*, 19, 711–26.
- Hays SM, Aylward LL. (2009). Using biomonitoring equivalents to interpret human biomonitoring data in a public health risk context. *J Appl Toxicol*, 29, 275–88.
- Hays SM, Aylward LL, LaKind JS, et al. (2008). Guidelines for the derivation of biomonitoring equivalents: report from the Biomonitoring Equivalents Expert Workshop. *Regul Toxicol Pharmacol*, 51, S4–15.
- Hays SM, Becker RA, Leung HW, et al. (2007). Biomonitoring equivalents: a screening approach for interpreting biomonitoring results from a public health risk perspective. *Regul Toxicol Pharmacol*, 47, 96–109.

- Hinderliter PM, Price PS, Bartels MJ, et al. (2011). Development of a source-to-outcome model for dietary exposures to insecticide residues: an example using chlorpyrifos. *Regul Toxicol Pharmacol*, 61, 82–92.
- IPCS (International Programme on Chemical Safety). (1994). Environmental Health Criteria 170: assessing human health risks of chemicals: derivation of guidance values for health-based exposure limits. Geneva, Switzerland.
- IPCS (International Programme on Chemical Safety). (2005). Chemical-specific adjustment factors for interspecies differences and human variability: guidance document for use of data in dose/concentration-response assessment. Geneva, Switzerland. Available from: [www.who.int/ipcs/methods/harmonization/areas/uncertainty/en/index.html](http://www.who.int/ipcs/methods/harmonization/areas/uncertainty/en/index.html) [last accessed 10 March 2013].
- IPCS (International Programme on Chemical Safety). (2006). IPCS framework for analysing the relevance of a cancer mode of action for humans and case studies. Geneva, Switzerland. Available from: [http://www.who.int/ipcs/methods/harmonization/areas/cancer\\_mode.pdf](http://www.who.int/ipcs/methods/harmonization/areas/cancer_mode.pdf) [last accessed 10 March 2013].
- IPCS (International Programme on Chemical Safety). (2010). Characterization and application of physiologically-based pharmacokinetic models in risk assessment. Available from: [http://www.who.int/ipcs/methods/harmonization/areas/pbpb\\_models.pdf](http://www.who.int/ipcs/methods/harmonization/areas/pbpb_models.pdf) [last accessed 10 March 2013].
- Jarabek AM. (1994). Inhalation RfC methodology: dosimetric adjustments and dose-response estimation of noncancer toxicity in the upper respiratory tract. *Inhal Toxicol*, 6, 301–25.
- Jarabek AM. (1995a). Interspecies extrapolation based on mechanistic determinants of chemical disposition. *J Hum Ecol Risk Assess*, 15, 641–52.
- 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, Menach MG, Overton JH, et al. (1989). Inhalation reference dose (RfDi): an application of interspecies dosimetry modeling for risk assessment of insoluble particles. *Health Phys*, 57, 177–83.
- Julien E, Boobis AR, Olin SS; The ILSI Research Foundation Threshold Working Group. (2009). The key events dose-response framework: a cross-disciplinary mode-of-action based approach to examining dose-response and thresholds. *Crit Rev Food Sci Nutr*, 49, 682–9.
- Kalberlah F, Schneider K. (1998). Quantification of extrapolation factors. Final report of the research project No. 116 06 113. Germany. Federal Environmental Agency.
- Keller DA, Juberg DR, Catlin N, et al. (2012). Identification and characterization of adverse effects in 21st century toxicology. *Toxicol Sci*, 126, 291–7.
- Kirman CR, Albertini RA, Gargas ML. (2010). 1,3-Butadiene: III. Assessing carcinogenic modes of action. *Crit Rev Toxicol*, 40, 74–92.
- Klassen C. (2008). Casarett & Doull's toxicology: the basic science of poisons, 7th ed. New York, NY: McGraw-hill medical Publishing Division. p. 23.
- Kroes R, Munro I, Poulsen E. (1993). Workshop on the scientific evaluation of the safety factor for the acceptable daily intake (ADI): editorial summary. *Food Add Contam*, 10, 269–73.
- LaKind JS, Aylward LL, Brunk C, et al. (2008a). Guidelines for the communication of biomonitoring equivalents: report from the biomonitoring equivalents expert workshop. *Regul Toxicol Pharmacol*, 51, 516–26.
- LaKind JS, Baraj L, Tran N, Aylward LL. (2008b). Environmental chemicals in people: challenges in interpreting biomonitoring information. *J Environ Health*, 70, 61–4.
- LaKind JS, Goodman M, Naiman DQ. (2012). Use of NHANES data to link chemical exposures to chronic diseases: a cautionary tale. *PLoS ONE*, 7, e51086. doi:10.1371/journal.pone.0051086
- Lau C, Andersen ME, Crawford-Brown DJ, et al. (2000). Evaluation of biologically based dose-response modeling for developmental toxicity: a workshop report. *Reg Appl Toxicol*, 31, 190–9.
- Lewis SC. (1993). Reducing uncertainty with adjustment factors. In: improvements in quantitative noncancer risk assessment. *Fund Appl Toxicol*, 20, 2–4.
- Lewis SC, Lynch JR, Nikiforov AI. (1990). A new approach to deriving community exposure guidelines from no-observed-adverse-effect levels. *Regul Toxicol Pharmacol*, 11, 314–30.
- Liao KH, Tan YM, Clewell III HJ. (2007). Development of a screening approach to interpret human biomonitoring data on volatile organic compounds: reverse dosimetry on biomonitoring data for trichloroethylene. *Risk Anal*, 27, 1223–36.
- Lipscomb JC, Haddad S, Poet T, Krishnan K. (2012). Physiologically-based pharmacokinetic (PBPK) models in toxicity testing and risk assessment. *Adv Exp Med Biol*, 745, 76–95.
- Lu FC. (1988). Acceptable daily intake: inception, evolution, and application. *Regul Toxicol Pharmacol*, 8, 45–60.
- MAK Commission. (2012). List of MAK and BAT Values 2012. Commission for the Investigation of Health Hazards of Chemical Compounds in the Work Area. Report No. 48. Deutsche Forschungsgemeinschaft. Wiley-VCH Verlag GmbH & Co. KGaA, pp. 158–9.
- Martin MT, Judson RS, Reif DM, et al. (2009). Profiling chemicals based on chronic toxicity results from the U.S. EPA ToxRef Database. *Environ Health Perspect*, 117, 392–9.
- Martin MT, Knudsen TB, Reif DM, et al. (2011). Predictive model of rat reproductive toxicity from ToxCast high throughput screening. *Biol Reprod*, 85, 327–39.
- McClellan RO. (2011). Role of science and judgment in setting national ambient air quality standards: how low is low enough? *Air Qual Atmos Health*, 5, 243–58. DOI 10.1007/s11869-011-0147-2.
- McLanahan ED, El-Masri HA, Sweeney LM, et al. (2012). Physiologically based pharmacokinetic model use in risk assessment—why being published is not enough. *Toxicol Sci*, 126, 5–15.
- Meek ME. (2008). Recent developments in frameworks to consider human relevance of hypothesized modes of action for tumours in animals. *Environ Mol Mutagen*, 49, 110–16.
- Meek ME, Bolger M, Bus JS, et al. (2013). A framework for “fit for purpose” dose response-assessment. *Regul Toxicol Pharmacol*, 66, 234–40.
- Meek ME, Boobis AR, Crofton KM, et al. (2011). Risk assessment of combined exposure to multiple chemicals: a WHO/IPCS framework. *Regul Toxicol Pharmacol*, 60, S1–14.
- Meek ME, Bucher JR, Cohen SM, et al. (2003). A framework for human relevance analysis of information on carcinogenic modes of action. *Crit Rev Toxicol*, 33, 591–653.
- Meek ME, Klaunig JE. (2010). Proposed mode of action of benzene-induced leukemia: interpreting available data and identifying critical data gaps for risk assessment. *Chem Biol Interact*, 184, 279–85.
- Moolgavkar SH, Knudson Jr AG. (1981). Mutation and cancer: a model for human carcinogenesis. *J Natl Cancer Inst*, 66, 1037–52.
- Meek ME, Newhook R, Liteplo RG, Armstrong VC. (1994). Approach to assessment of risk to human health for priority substances under the Canadian Environmental Protection Act. *Environ Carcin Eco R*, C12, 105–34.
- Naumann B, Meek ME, Dourson, ML, Ohanian, E. (2005). The future of chemical specific adjustment factors in risk assessment. *Risk Policy Report*, 12, 14. Available from: <http://www.tera.org/Publications/risk%20policy%20report-nauman%20et%20al%202005.pdf> [last accessed 10 March 2013].
- NRC (National Research Council). (1983). Risk assessment on the federal government-managing the process. Washington, DC: National Academies of Science, National Academy Press.
- NRC (National Research Council). (1989). Recommended daily allowances. Washington, DC: National Academy of Science, National Academy Press.
- NRC (National Research Council). (1993). Issues in risk assessment. Washington, DC: National Academies of Science, National Academy Press.
- NRC (National Research Council). (1994). Science and judgment in risk assessment. Washington, DC: National Academies of Science, National Academy Press.
- NRC (National Research Council). (1996). Understanding risk. Washington, DC: National Academies of Science, National Academy Press.
- NRC (National Research Council). (2004). Review of the army's technical guides on assessing and managing chemical hazards to deployed personnel. Washington, DC: National Academies of Science, National Academy Press.
- NRC (National Research Council). (2006). Human Biomonitoring for Environmental Chemicals. Committee on Human Biomonitoring for Environmental Toxicants. Washington, DC: National Academies of Science, National Academy Press.

- NRC (National Research Council). (2007a). Toxicity testing in the 21st century: a vision and a strategy. Washington, DC: National Academies of Science, National Academy Press.
- NRC (National Research Council). (2007b). Applications of toxicogenomic technologies to predictive toxicology and risk assessment. Washington, DC: National Academy Press.
- NRC (National Research Council). (2008a). Public participation in environmental assessment and decision making. Washington, DC: National Academies of Science, National Academy Press.
- NRC (National Research Council). (2008b). Phthalates and cumulative risk assessment: the tasks ahead. Washington, DC: National Academies of Science, National Academy Press.
- NRC (National Research Council). (2009). Science and decisions: advancing risk assessment. Washington, DC: National Academies of Science, National Academy Press.
- OECD (Organisation for Economic Co-operation and Development). (2007). Guidance document on the validation of (quantitative) structure-activity relationship [(Q)SAR] models. OECD Environment Health and Safety Publications Series on Testing and Assessment, No. 69. ENV/JM/MONO(2007)2.
- Olin SS, Sonawane BR. (2003). Workshop to develop a framework for assessing risks to children from exposure to environmental agents. *Environ Health Persp*, 111, 1524–26.
- Olipitz W, Wiktor-Brown D, Shuga J, et al. (2012). Integrated molecular analysis indicates undetectable change in DNA damage in mice after continuous irradiation at ~400-fold natural background radiation. *Environ Health Perspect*, 120, 1130–36.
- Pieters MN, Kramer HJ, Slob W. (1998). Evaluation of the uncertainty factor for subchronic-to-chronic extrapolation: statistical analysis of toxicity data. *Regul Toxicol Pharmacol*, 27, 108–18.
- Pottenger LH, Gollapudi BB. (2010). Genotoxicity testing: moving beyond qualitative “screen and bin” approach towards characterization of dose-response and thresholds. *Environ Mol Mutagen*, 51, 792–9.
- Pottenger LH, Schisler MR, Zhang F, et al. (2009). Dose-response and operational thresholds/NOAELs for *in vitro* mutagenic effects from DNA-reactive mutagens, MMS and MNU. *Mutat Res*, 678, 138–47.
- President's Commission. (1997). Presidential/Congressional Commission on Risk Assessment and Risk Management. Available from: <http://www.riskworld.com/Nreports/1997/risk-rpt/pdf/EPAJAN.PDF> [last accessed 10 March 2013].
- Price PS, Han X. (2011). Maximum cumulative ratio (MCR) as a tool for assessing the value of performing a cumulative risk assessment. *Int J Environ Res Public Health*, 8, 2212–25.
- Price PS, Schnelle KD, Cleveland CB, et al. (2011). Application of a source-to-outcome model for the assessment of health impacts from dietary exposures to insecticide residues. *Regul Toxicol Pharmacol*, 61, 23–31.
- Renwick AG. (1991). Safety factors and establishment of acceptable daily intake. *Food Addit Contam*, 8, 135–50.
- Renwick AG. (1993). Data derived safety factors for the evaluation of food additives and environmental contaminants. *Food Addit Contam*, 10, 275–305.
- Renwick AG. (1998a). Toxicokinetics in infants and children in relation to the ADI and TDI. *Food Addit Contam*, 15, 17–35.
- Renwick AG, Lazarus NR. (1998b). Human variability and noncancer risk assessment—an analysis of the default uncertainty factor. *Regul Toxicol Pharmacol*, 27, 3–20.
- Renwick AG, Dorne JL, Walton K. (2000). An analysis of the need for an additional uncertainty factor for infants and children. *Regul Toxicol Pharmacol*, 31, 286–96.
- Renwick AG, Dorne JCM, Walton K. (2001). Pathway-related factors: the potential for human data to improve the scientific basis of risk assessment. *HERA*, 7, 165–80.
- Renwick AG, Barlow SM, Hertz-Picciotto I, et al. (2003). Risk characterisation of chemicals in food and diet. *Food Chem Toxicol*, 41, 1211–71.
- Rhomberg LR. (2011). Alliance for risk assessment (ARA) presentation. Available from: [http://www.allianceforrisk.org/ARA\\_Dose-Response.htm](http://www.allianceforrisk.org/ARA_Dose-Response.htm) [last accessed 10 March 2013].
- Rhomberg LR, Goodman JE, Haber LT, et al. (2011). Linear low-dose extrapolation for noncancer health effects is the exception, not the rule. *Crit Rev Toxicol*, 41, 1–19.
- Sand S, Portier CJ, Krewski D. (2011). A signal-to-noise crossover dose as the point of departure for health risk assessment. *Environ Health Perspect*, 119, 1766–74.
- Schulte PA. (1989). A conceptual framework for the validation and use of biological markers. *Environ Res*, 48, 129–44.
- SCOEL. (2010). Recommendation from the scientific committee on occupational exposure limits for propylene oxide. SCOEL/SUM/161. August 2010. Available from: <http://ec.europa.eu/social/BlobServlet?docId=6514&langId=en> [last accessed 10 Mar 2013].
- Seed J, Carney EW, Corley RA, et al. (2005). Overview: using mode of action and life stage information to evaluate the human relevance of animal toxicity data. *Crit Rev Toxicol*, 35, 664–72.
- Shah I, Houck K, Judson RS, et al. (2011). Using nuclear receptor activity to stratify hepatocarcinogens. *PLoS ONE*, 6, e14584. doi:10.1371/journal.pone.0014584.
- Shuey DL, Setzer RW, Lau C, et al. (1995). Biological modeling of 5-fluorouracil developmental toxicity. *Toxicology*, 102, 207–13.
- Silverman KC, Naumann BD, Holder DJ, et al. (1999). Establishing data-derived adjustment factors from published pharmaceutical clinical trial data. *HERA*, 5, 1059–89.
- Sipes NS, Martin MT, Reif DM, et al. (2011). Predictive models of prenatal developmental toxicity from ToxCast high-throughput screening data. *Toxicol Sci*, 124, 109–27.
- Slikker Jr W, Andersen ME, Bogdanffy MS, et al. (2004a). Dose-dependent transitions in mechanisms of toxicity. *Toxicol Appl Pharm*, 201, 203–25.
- Slikker Jr W, Andersen ME, Bogdanffy MS, et al. (2004b). Dose-dependent transitions in mechanisms of toxicity: case studies. *Toxicol Appl Pharm*, 201, 226–94.
- Sonich-Mullin C, Fielder R, Wiltse J, et al. (2001). International programme on chemical safety. IPCS conceptual framework for evaluating a mode of action for chemical carcinogenesis. *Regul Toxicol Pharmacol*, 34, 146–52.
- Souza T, Kush R, Evans JP. (2007). Global clinical data exchange standards are here! *Drug Discov Today*, 12, 174–81.
- Strawson J, Haber L, Dourson M, Barkhurst M. (2003). Approaches to determining “Unreasonable Risk to Health.” Toxicology Excellence for Risk Assessment (TERA). December 29. Available from: <http://nrwa.org/search/mainresultland.aspx?cx=partner-pub-1008341574573985%3A0126459385&cof=FORID%3A10&ie=UTF-8&q=TERA> [last accessed 10 March 2013].
- Suter GW, Vermeire T, Munns Jr WR, Sekizawa J. (2003). Framework for the integration of health and ecological risk assessment. *Hum Ecol Risk Assess*, 9, 281–301.
- Swartout JC, Price PS, Dourson ML, et al. (1998). A probabilistic framework for the reference dose (probabilistic RfD). *Risk Anal*, 18, 271–82.
- Sweeney LM, Kirman CR, Albertini RJ, et al. (2009). Derivation of inhalation toxicity reference values for propylene oxide using mode of action analysis: example of a threshold carcinogen. *Crit Rev Toxicol*, 39, 462–86.
- Swenberg JA, Richardson FC, Boucheron JA, et al. (1987). High- to low-dose extrapolation: critical determinants involved in the dose response of carcinogenic substances. *Environ Health Persp*, 76, 57–63.
- Swenberg JA, Lu K, Moeller BC, et al. (2011). Endogenous versus exogenous DNA adducts: their role in carcinogenesis, epidemiology, and risk assessment. *Toxicol Sci*, 120, S130–45.
- Teuschler LK, Dourson ML, Stiteler WM, et al. (1999). Health risk above the reference dose for multiple chemicals. *Reg Toxicol Pharmacol*, 30, S19–26. Available from: <http://www.tera.org/Publications/Teuschler%20et%20al%201999.pdf> [last accessed 10 March 2013].
- Thomas RS, Black MB, Li L, et al. (2012a). A comprehensive statistical analysis of predicting *in vivo* hazard using high-throughput *in vitro* screening. *Toxicol Sci*, 128, 398–417.
- Thomas RS, Clewell III HJ, Allen BC, et al. (2012b). Integrating pathway-based transcriptomic data into quantitative chemical risk assessment: a five chemical case study. *Mutat Res*, 746, 135–43.
- Truhaut R. (1991). The concept of acceptable daily intake: a historical review. *Food Addit Contam*, 8, 151–62.
- USDA (U.S. Department of Agriculture). (2012). Microbial risk assessment guideline. Pathogenic microorganisms with focus on food and water. Interagency Microbiological Risk Assessment

- Guideline Workgroup July 2012. Publication Numbers: USDA/FSIS/2012-001 EPA/100/J12/00.
- US EPA (U.S. Environmental Protection Agency). (1976). Interim Procedures and Guidelines for Health Risk and Economic Impact Assessments of Suspected Carcinogens. Available from: [nepis.epa.gov/Exe/ZyPURL.cgi?Dockey=90160G00.txt](http://nepis.epa.gov/Exe/ZyPURL.cgi?Dockey=90160G00.txt) [last accessed 24 Feb 2013].
- US EPA (U.S. Environmental Protection Agency). (1986a). Guidelines for carcinogen risk assessment. Federal Register, 51, 33992–34003.
- US EPA (U.S. Environmental Protection Agency). (1986b). Guidelines for the health risk assessment of chemical mixtures. Federal Register, 51, 34014–25.
- US EPA (U.S. Environmental Protection Agency). (1989). Risk Assessment Guidance for Superfund. Volume I. Human Health Evaluation Manual (Part A). Washington, DC. EPA/540/1-89/002.
- US EPA (U.S. Environmental Protection Agency). (1992). Risk Assessment Forum. Framework for Ecological Risk Assessment. Washington, DC. EPA/630/R-92/001.
- US EPA (U.S. Environmental Protection Agency). (1994). Methods for Derivation of Inhalation Reference Concentrations and Application of Inhalation Dosimetry. Office of Health and Environmental Assessment. Washington, DC. EPA/600/8-90-066F.
- US EPA (U.S. Environmental Protection Agency). (1996). Proposed guidelines for carcinogen risk assessment. Federal Register, 61, 17960-18011.
- US EPA (U.S. Environmental Protection Agency). (1997). Office of Solid Waste and Emergency Response. Ecological Risk Assessment Guidance for Superfund: process for Designing and Conducting Ecological Risk Assessments - Interim Final. EPA 540-R-97-006. Available from: [http://www.epa.gov/oswer/riskassessment/superfund\\_problem.htm](http://www.epa.gov/oswer/riskassessment/superfund_problem.htm) [last accessed 12 Sep 2011].
- US EPA (U.S. Environmental Protection Agency). (1998). Risk Assessment Forum. Guidelines for Ecological Risk Assessment. Washington, DC. EPA/630/R-95/002F.
- US EPA (U.S. Environmental Protection Agency). (1999). Office of Solid Waste and Emergency Response. Ecological Risk Assessment and Risk Management Principles for Superfund Sites, (Issuance of Final Guidance) (PDF) October 7, 1999. Available from: [http://www.epa.gov/oswer/riskassessment/superfund\\_problem.htm](http://www.epa.gov/oswer/riskassessment/superfund_problem.htm) [last accessed 12 Sep 2011].
- US EPA (U.S. Environmental Protection Agency). (2000a). Handbook-Risk Characterization. Science Policy Council. Office of Science Policy. Office of Research and Development. EPA 100-B-00-002.
- US EPA (U.S. Environmental Protection Agency). (2000b). Supplementary Guidance for Conducting Health Risk Assessment of Chemical Mixtures. EPA/630/R-00/002.
- US EPA (U.S. Environmental Protection Agency). (2001). Office of Solid Waste and Emergency Response. Risk Assessment Guidance for Superfund, Volume I - Human Health Evaluation Manual. Part D: standardized Planning, Reporting and Review of Superfund Risk Assessments. Available from: [http://www.epa.gov/oswer/riskassessment/superfund\\_hhplanning.htm](http://www.epa.gov/oswer/riskassessment/superfund_hhplanning.htm) [last accessed 12 September 2011].
- US EPA (U.S. Environmental Protection Agency). (2002a). A review of the Reference Dose (RfD) and Reference Concentration (RfC) processes. Risk Assessment Forum. EPA/630/P-02/002F.
- US EPA (U.S. Environmental Protection Agency). (2002b, February 28). Determination of the appropriate FQPA safety factor(s) in tolerance assessment. Washington, DC: Office of Pesticide Programs.
- US EPA (U.S. Environmental Protection Agency). (2003a). Framework for Cumulative Risk Assessment. Risk Assessment Forum. U.S. Environmental Protection Agency. Washington, DC. EPA/630/P-02/001F.
- US EPA (U.S. Environmental Protection Agency). (2003b). Region 3. Multi-criteria Integrated Resource Assessment (MIRA). Available from: <http://www.epa.gov/reg3esd1/data/mira.htm> [last accessed 9 Feb 2012].
- US EPA (U.S. Environmental Protection Agency). (2005). Guidelines for carcinogen risk assessment. U.S. Environmental Protection Agency. EPA/630/P-03/001b.
- US EPA (U.S. Environmental Protection Agency). (2006a). A Framework for Assessing Health Risk of Environmental Exposures to Children (Final). National Center for Environmental Assessment. Office of Research and Development. Washington, DC, EPA/600/R-05/093F.
- US EPA (U.S. Environmental Protection Agency). (2006b). Office of Pesticide Programs. Registration Review Process. Available from: [http://www.epa.gov/pesticides/oppsrrd1/registration\\_review/reg\\_review\\_process.htm](http://www.epa.gov/pesticides/oppsrrd1/registration_review/reg_review_process.htm) [last accessed 10 Feb 2012].
- US EPA (U.S. Environmental Protection Agency). (2006c). Approaches for the Application of Physiologically Based Pharmacokinetic (PBPK) Models and Supporting Data in Risk Assessment (Final Report). National Center for Environmental Assessment, Washington, DC. EPA/600/R-05/043F.
- US EPA (U.S. Environmental Protection Agency). (2007). Framework for Metals Risk Assessment. Risk assessment Forum. Washington, DC. EPA 120/R-07/001.
- US EPA (U.S. Environmental Protection Agency). (2009). Office of Air Quality Planning and Standards. Process for Reviewing National Ambient Air Quality Standards. Available from: <http://www.epa.gov/ttn/naaqs/review2.html> [last accessed 9 Feb 2012].
- US EPA (U.S. Environmental Protection Agency). (2011a). U.S. Environmental Protection Agency & U.S. Department of Agriculture Tolerance Reassessment Advisory Committee of the National Advisory Council for Environmental Policy & Technology. Available from: <http://www.epa.gov/oppead1/trac/> [last accessed 12 Sep 2011].
- US EPA (U.S. Environmental Protection Agency). (2011b). U.S. Environmental Protection Agency & U.S. Department of Agriculture Committee to Advise on Reassessment and Transition (CARAT) National Advisory Council for Environmental Policy & Technology. Available from: <http://www.epa.gov/oppead1/carat/> [last accessed 12 Sep 2011].
- US EPA (U.S. Environmental Protection Agency). (2011c). Public Involvement in Registration Review. Available from: [http://www.epa.gov/oppsrrd1/registration\\_review/public\\_involvement.htm](http://www.epa.gov/oppsrrd1/registration_review/public_involvement.htm) [last accessed 12 Sep 2011].
- US EPA (U.S. Environmental Protection Agency). (2011d). Office of Pesticide Programs. Ecological Risk Assessment: technical Overview. Available from: [http://www.epa.gov/oppead1/ecorisk\\_ders/](http://www.epa.gov/oppead1/ecorisk_ders/) [last accessed 12 Sep 2011].
- US EPA (U.S. Environmental Protection Agency). (2011e). Guidance for applying quantitative data to develop data-derive extrapolation factors for interspecies and intraspecies extrapolation. External Review Draft, May. Risk Assessment Forum. EPA/100/J-11/001. Available from: <http://www.epa.gov/raf/files/ddef-external-review-draft05-11-11.pdf> [last accessed 1 Aug 2012].
- US EPA (U.S. Environmental Protection Agency). (2012a). Framework for Human Health Risk Assessment to Inform Decision Making. EPA External Review Draft. Office of the Science Advisor. Risk Assessment Forum. Washington, DC. 601-D12-001. Available from: <http://www.epa.gov/raf/files/framework-document-7-13-12.pdf> [last accessed 1 Aug 2012].
- US EPA (U.S. Environmental Protection Agency). (2012b). Office of Pesticide Programs. Pesticide Registration Review Status. Available from: [http://www.epa.gov/oppsrrd1/registration\\_review/reg\\_review\\_status.htm](http://www.epa.gov/oppsrrd1/registration_review/reg_review_status.htm) [last accessed 9 Feb 2012].
- US EPA (U.S. Environmental Protection Agency). (2012c). Environmental Protection Agency: notice of workshop and call for information on integrated science assessment for oxides of nitrogen. Federal Register, 77, 7149-7151.
- US EPA (U.S. Environmental Protection Agency). (2012d). Office of Water. Integrated Planning Approach Framework (draft). Available from: [http://www.epa.gov/npdes/pubs/integrated\\_planning\\_framework\\_draft.pdf](http://www.epa.gov/npdes/pubs/integrated_planning_framework_draft.pdf) [last accessed 9 Feb 2012].
- US EPA (U.S. Environmental Protection Agency). (2012e). Vocabulary Catalog List Detail - Integrated Risk Information System (IRIS) Glossary. Available from: [http://ofmpub.epa.gov/sor\\_internet/registry/termreg/searchandretrieve/glossariesandkeywordlists/search.do?details=&glossaryName=IRIS%20Glossary](http://ofmpub.epa.gov/sor_internet/registry/termreg/searchandretrieve/glossariesandkeywordlists/search.do?details=&glossaryName=IRIS%20Glossary). [last accessed 1 Aug 2012].
- US EPA (U.S. Environmental Protection Agency). (2012f). Integrated Risk Information System (IRIS) Chloroform. Available from: <http://www.epa.gov/iris> [last accessed 10 March 2013].
- US EPA (U.S. Environmental Protection Agency). (2012g). Advances in Inhalation Gas Dosimetry for Derivation of a Reference Concentration (RfC) and Use in Risk Assessment. EPA/600/R-12/044.
- Vines TH, Andrew RL, Bock DG, et al. (2013). Mandated data archiving greatly improves access to research data. *FASEB J*, 27, 1304–8.

- Waters MD, Fostel JM. (2004). Toxicogenomics and systems toxicology: aims and prospects. *Nat Rev Genet*, 5, 936–48.
- WHO IPCS (World Health Organization. International Programme on Chemical Safety). (2006). Environmental Health Criteria 237-Principles for Evaluating Health Risks in Children Associated with Exposure to Chemicals. Geneva, Switzerland.
- WHO IPCS (World Health Organization. International Programme on Chemical Safety). (2007). PART 1: IPCS Framework for Analysing The Relevance of a Cancer Mode of Action for Humans And Case-Studies. PART 2: IPCS Framework for Analysing the Relevance of a Non-Cancer Mode of Action for Humans. Geneva, Switzerland.
- WHO IPCS (World Health Organization. International Programme on Chemical Safety). (2008). Uncertainty and data quality in exposure assessment. Part 1. Guidance document on characterizing and communicating uncertainty in exposure assessment. Harmonization Project Document No. 6. Geneva, Switzerland.
- WHO IPCS (World Health Organization. International Programme on Chemical Safety). (2009a). Environmental Health Criteria 239 - Principles for Modeling Dose-Response for the Risk Assessment of Chemicals. Geneva, Switzerland.
- WHO IPCS (World Health Organization. International Programme on Chemical Safety). (2009b). Environmental Health Criteria 240 - Principles and methods for the risk assessment of chemicals in food. Geneva, Switzerland.
- WHO IPCS (World Health Organization. International Programme on Chemical Safety). (2010). Characterization and Application of Physiologically Based Pharmacokinetic Models in Risk Assessment. Harmonization Project Document No. 9. Geneva, Switzerland.
- WHO (World Health Organization). (2011). Guidelines for Drinking-water Quality. Fourth Edition. Geneva, Switzerland.
- Whyatt RM, Rauh V, Barr DB, et al. (2004). Prenatal insecticide exposures and birth weight and length among an urban minority cohort. *Environ Health Perspect*, 112, 1125–32.
- Woodruff TJ, Zota AR, Schwartz JM. (2011). Environmental Chemicals in Pregnant Women in the United States: NHANES 2003–2004. *Environ Health Perspect*, 119, 878–885. Available from: <http://ehp03.niehs.nih.gov/article/info:doi/10.1289/ehp.1002727> [last accessed 10 March 2013].
- Yang Y, Abel SJ, Ciurlionis R, Waring JF. (2006). Development of a toxicogenomics in vitro assay for the efficient characterization of compounds. *Pharmacogenomics*, 7, 177–86.
- Zhao Q, Gadagbui B, Dourson M. (2005). Lower birth weight as a critical effect of Chlorpyrifos: a comparison of human and animal data. *Reg Toxicol Pharmacol*, 42, 55–63.
- Zhao Q, Unrine J, Dourson ML. (1999). Replacing the default values of 10 with data-derived values: a comparison of two different data-derived uncertainty factors for boron. *Hum Ecol Risk Assess*, 5, 973–83. Available from: <http://www.tera.org/Publications/boron.pdf> [last accessed 10 March 2013].
- Zielhuis RL, van der Kreek FW. (1979). The use of a safety factor in setting health based permissible levels for occupational exposure. *Int Arch Occup Environ Health*, 42, 191–201.