



# ENVIRONMENTAL HEALTH PERSPECTIVES

<http://www.ehponline.org>

## Key Elements for Judging the Quality of a Risk Assessment

Penelope A. Fenner-Crisp and Vicki L. Dellarco

<http://dx.doi.org/10.1289/ehp.1510483>

Received: 14 July 2015

Accepted: 21 January 2016

Advance Publication: 5 February 2016

**Note to readers with disabilities:** *EHP* will provide a [508-conformant](#) version of this article upon final publication. If you require a 508-conformant version before then, please contact [ehp508@niehs.nih.gov](mailto:ehp508@niehs.nih.gov). Our staff will work with you to assess and meet your accessibility needs within 3 working days.



National Institute of  
Environmental Health Sciences

## Key Elements for Judging the Quality of a Risk Assessment

Penelope A. Fenner-Crisp<sup>1</sup> and Vicki L. Dellarco<sup>2</sup>

<sup>1</sup>Independent Consultant, North Garden, Virginia, USA; <sup>2</sup>Independent Consultant, Silver Spring, Maryland, USA. (Note: Both authors are retired from the U.S. Environmental Protection Agency)

**Address correspondence to** Penelope A. Fenner-Crisp, 2871 Southern Hills Drive, North Garden, VA 22959 USA. Telephone: 434-244-2691. E-mail: [pfennercrisp@aol.com](mailto:pfennercrisp@aol.com)

**Running title:** Guide for judging quality of risk assessments

**Competing financial interests:** Preparation of this manuscript was funded in part by the American Chemistry Council (ACC).

The authors have sole responsibility for the writing and contents of this paper.

This paper represents the professional opinions of the authors and not necessarily those of the ACC.

## **Abstract**

**Background:** Many reports have been published which contain recommendations for improving the quality, transparency and usefulness for decision-making of risk assessments prepared by agencies of the U.S. federal government. A substantial measure of consensus has emerged as to what characteristics high quality assessments should possess.

**Objective:** The goal was to summarize the key characteristics of a high quality assessment as identified in the consensus-building process and integrate them into a Guide for use by decision-makers, risk assessors, peer reviewers and other interested stakeholders to determine if an assessment meets the criteria for high quality.

**Discussion:** Most of the features cited in the Guide are applicable to any type of assessment, whether it encompasses just one, or two, or all four phases of the risk assessment paradigm, whether qualitative or quantitative, screening level or highly sophisticated and complex. Other features are tailored to specific elements of an assessment. Just as agencies at all levels of government are responsible for determining the effectiveness of their programs, so should they determine the effectiveness of their assessments used in support of their regulatory decisions. Furthermore, if a non-governmental entity wishes to have its assessments considered in the governmental regulatory decision-making process, then these should be judged in the same rigorous manner and held to similar standards.

**Conclusions:** The key characteristics of a high quality assessment can be summarized and integrated into a Guide for judging whether an assessment possesses the desired features of high quality, transparency and usefulness.

## **Introduction**

A number of U.S. federal (as well as state and local) government agencies produce risk assessments on a continuing basis. In the years since publication of the 1983 National Research Council report *Risk Assessment in the Federal Government: Managing the Process* (the “Red Book,” NRC 1983), advances in risk assessment have occurred but the need for further improvement continues to be recognized. Much attention has been focused on the assessment practices of the U.S. Environmental Protection Agency (US EPA), although recommendations for improvement have been directed toward other agencies, as well. In our opinion, the problems ascribed by critics to these assessments generally do not lie in the lack of guidance on how to conduct an assessment, but rather in the failure to implement internal guidance or externally-generated recommendations in a consistent and transparent manner. The aim of this paper is to extract from the accumulated recommendations of many expert panels a set of attributes that can serve as a guide for judging whether or not an assessment has incorporated consensus best practices that result in a scientifically credible, transparent and useful product. By “best practices,” we mean that an assessment possesses scientific accountability and integrity by employing a critical, open-minded approach in selecting reliable data and models fit for their intended use, and analyzing and integrating that information. It uses defined methodologies for collecting and interpreting information and minimizing any bias that might be introduced. Its development process embraces the necessary scoping and planning before conducting the assessment. It ensures that transparency exists throughout to enable others to judge the scientific robustness of the conclusions and to replicate the findings and it describes the uncertainties associated with the assessment. And, finally, it is readily-usable and provides value to its intended audiences.

## **Objective**

The Guide presented in Appendix 1 has been designed for use by decision-makers to assist in their quest to have a high quality assessment at hand when carrying out their responsibilities and by authors, sponsors, risk assessors, peer reviewers and other interested stakeholders to determine if an assessment “measures up” to current best scientific practices. The use of the Guide is intended to promote transparency and consistency with regards to the conduct and quality of assessments.

## **Methods**

A general consensus has been evolving over the past several years as to what characteristics high quality assessments should possess. A review of a series of primarily government-funded expert panel reports was conducted in order to identify, assemble and synthesize the key elements of a high quality assessment for the purpose of creating a simple and useful quality assurance guide. These reports included those of the National Research Council (NRC), the Institute of Medicine (IOM), the Presidential/Congressional Commission on Risk Assessment and Risk Management and a number of foreign governments and organizations (*e.g.*, NRC 1983, 1994, 1996, 2007, 2009, 2011, 2013, 2014; IOM 2011; Health Canada 2000, 2015; EFSA 2010, 2011, 2014a, b; ECHA 2011; OECD 2007, 2012).

## **Reports Relevant to the Development of the Guide for Judging the Quality of an Assessment**

The processes by which risk assessments are developed as well as their substance and content have been the subject of deliberation by many parties over the last 30+ years. The U.S. Congress, the Executive Branch, various commissions, National Research Council (NRC) and Institute of Medicine (IOM) committees, affected stakeholder communities, even the general

public and individuals have all weighed in. Over time, there has been a shift in, and an expansion of, the areas of focus on the elements of the risk assessment process. Perhaps, this evolution can best be illustrated by tracking the topics addressed in a series of reports, primarily from the National Research Council, that began with the publication of the 1983 Red Book. Congress directed the U.S. Food and Drug Administration (FDA) to contract with the National Academy of Sciences (NAS) to conduct a study of the institutional means for risk assessment (NRC, 1983). Of particular interest at that time was the interface and interaction between science and policy, and between risk assessment and risk management. The focus was on the potential for human health impacts of exposure to chemicals. In reality, however, the discussion is equally applicable to ecological risk assessment, to other stressors such as radiation, microbes and products of biotechnology, and to many categories such as environmental contaminants, food additives, constituents and contaminants, medical devices, drugs, tobacco, consumer products, commodity chemicals and pesticides.

The Red Book committee made several recommendations for improving risk assessment through changes in procedures such as 1) maintaining a clear distinction between the science and the other factors involved in decision-making, including political considerations, economics, and technology; 2) making a risk assessment document publically-available before finalizing regulatory decisions; 3) subjecting the risk assessment to external expert peer review, and 4) developing joint assessments if two or more agencies have a regulatory mandate regarding the same chemical(s) (NRC 1983). This committee also offered recommendations on improving risk assessment through the development of uniform assessment guidelines. In its view, there would be one set of guidelines that all agencies would implement. These guidelines would be detailed, but flexible, and address all four phases of risk assessment: hazard identification, dose-

response assessment, exposure assessment and risk characterization. Guidelines for assessing cancer risk would be developed first, then for other endpoints of toxicity and for exposure. Furthermore, they would be developed by a congressionally-chartered board of experts who were independent of regulatory decision-making.

Although it was FDA that funded this report, it was EPA under the leadership of Administrator William Ruckelshaus that most vigorously embraced the recommendations, implementing many of them (US EPA 1993). EPA implemented all of the recommendations on procedural changes, although there are only a few examples of collaboration with other federal agencies on specific chemicals. From the mid-1980s through the 1990s, EPA developed guidelines for cancer, mutagenicity, reproductive and developmental toxicity, neurotoxicity, chemical mixtures and ecological effects to promote consistency agency-wide. Many other EPA-wide policies, principles and risk assessment guidance, databases, models and other tools have been developed since then. Efforts to develop guidance for inter-agency use have not succeeded.

The Red Book, then, serves as the starting point for discussion. A number of the observations and recommendations in subsequent reports can be traced back to concepts originally articulated in it. The 1994 NRC report *Science and Judgment in Risk Assessment* emphasized approaches to exposure and toxicity assessment and risk characterization as well as strategies for improving risk assessment in the areas of default options, models, data needs, uncertainty, variability and aggregation of risk (NRC1994). A brief discussion of priority-setting laid the groundwork for future findings and recommendations related to the use of pre-planning and problem formulation measures before a resource-intensive assessment is begun. An expanded discussion of problem formulation as well as the desirability and importance of expert scientific peer review and for

input and comment from interested stakeholders outside of the organization which prepared the assessment was addressed in the 1996 NRC report *Understanding Risk: Informing Decisions in a Democratic Society* (NRC 1996). The value of problem formulation and planning and scoping prior to conducting an assessment also was emphasized in the findings and recommendations of the Presidential/Congressional Commission on Risk Assessment and Risk Management when it stated that “The level of detail considered in a risk assessment and included in a risk characterization should be commensurate with the problem’s importance, expected health or environmental impact, expected economic or social impact, urgency, and level of controversy, as well as with the expected impact and cost of protective measures” (Presidential/Congressional Commission 1997).

The 2009 NRC report *Science and Decisions: Advancing Risk Assessment* revisited topics such as uncertainty and variability, defaults and cumulative risk assessment along with new issues related to dose response assessment (NRC 2009). The importance of the design of risk assessment processes to improve the utility for decision-making also was addressed. This committee observed that “the selection of appropriate elements of process and the specification of required elements of the final product constitute a complex design challenge.” They viewed the incorporation of “fairness, transparency and efficiency” in both the process and the resulting assessments as critical elements in assuring the quality and usefulness of the assessments, both to decision-makers and to other stakeholders. “Objectivity” and “balance” also are essential characteristics of high quality products. To provide some structure to the risk assessment process, this committee presented its vision of a framework for risk-based decision-making that, in its view, would make the best use of a risk assessment. Similar in structure and content to the framework schematics included in a number of EPA’s guidance documents (*e.g.*, US EPA

1992, 2003, 2006, 2014a), it describes three general components, the first including problem formulation and scoping, the second reflecting planning as well as the technical components of the risk assessment itself, and the third focused on the other factors (*e.g.*, legal, technological, economic) that must be considered to reach and communicate management decisions. Risk assessment frameworks developed by other governments and organizations also emphasize the importance of problem formulation as a first step (*e.g.*, Health Canada 2000; WHO 2010; EFSA 2013).

Continuing dissatisfaction with the perceived lack of adequate documentation and transparency in the selection and interpretation of data and the application of science policy guidance as well as the perception that many risk assessments do not reflect the incorporation of the best available science has focused attention on the concept of systematic review (NRC 2014). Systematic review is defined by the Institute of Medicine (IOM) as “a scientific investigation that focuses on a specific question and uses explicit, pre-specified scientific methods to identify, select, assess, and summarize the findings of similar but separate studies”(IOM 2011). It is contended that implementation of such an approach would serve to ameliorate at least some of the concerns.

Systematic review has been used for several decades in the fields of medicine, education and agriculture. If done properly, it can improve the credibility of the assessment. Although the early focus of systematic review was on clinical medicine and health care, beginning with the development of the Cochrane Collaboration (Higgins and Green 2011), the importance of a more formalized procedure is now seen to be important for human health and ecological assessments of chemical and environmental exposures (*e.g.*, US EPA 2013; TCEQ 2014; Woodruff and Sutton 2014; NTP 2015) food and feed safety (*e.g.*, US FDA 2009; EFSA 2010) and guideline

development (WHO 2012). In its report *Review of the Environmental Protection Agency's Draft IRIS Assessment of Formaldehyde*, the NRC argued strongly for the implementation of a systematic review process in the development of Integrated Risk Information System (IRIS) assessments (NRC 2011). While acknowledging that EPA had made some progress on incorporating elements of systematic review into its IRIS document development process since 2011, the NRC Committee to Review the IRIS Process pressed forward with additional comments and recommendations on problem formulation and protocol development, evidence identification, evaluation and integration for hazard characterization as well as methodological issues related to dose response assessment and the derivation of toxicity values (NRC 2014). This committee concluded that the general approaches and concepts underlying systematic reviews for evidence-based medicine embodied in the standards established by IOM should generally be relevant to the review of animal, epidemiologic and mechanistic studies in the IRIS hazard characterization process. One might argue that it also would be relevant to assessments prepared by other parties as well. Systematic review also should be presumed to be applicable and useful in those assessments that include exposure assessment and risk characterization.

Systematic review already has been embraced by the National Toxicology Program (NTP 2015) and the European Food Safety Authority (EFSA 2010, 2011) and is beginning to being implemented by several EPA Offices (US EPA 2012, 2013). While some efforts focus on hazard identification alone (NTP 2015) or on hazard identification and dose-response assessment (EFSA 2010; US EPA 2013), others cover exposure assessment and risk characterization as well (EFSA 2011; US EPA 2012).

The later steps of the systematic review process--interpreting results and drawing conclusions--flow into the assessment itself. The assessment is conducted against the backdrop of a pre-determined scope that defines the linkages between stressors (chemical or other) and adverse human health or ecological effects, including identifying the stressor(s), exposure pathway(s), exposed life stage(s) and population(s), and toxicity endpoint(s) of concern that will be addressed in the assessment (US EPA 1992, 1998, 2014a). The result of this effort is a completed assessment, constituting several components, which may or may not be issued at the same time. They might be staged, beginning with a problem formulation/planning and scoping product issued first, perhaps being subjected to peer review and public comment. A second product, the systematic review, also might be subjected to peer review and public comment. And, finally, the assessment itself—reflecting an objective, scientific analysis of the key data with a transparent identification of relevant science policy choices (*e.g.*, application of defaults, selection of dose response models, use of uncertainty factors) to be subjected to peer review and, perhaps, public comment.

### **Discussion of the Guide for Judging the Quality of an Assessment**

The ultimate purpose of the Guide is to provide guidance for evaluating the quality of an assessment. We envision the Guide to be used both as a self-assessment tool by the author(s) of an assessment and as a mechanism for judging the quality of an assessment prepared by another party. For purposes of discussion and simplification, we are viewing transparency and usefulness as desirable characteristics of quality and are folding them into the single term of “quality.” How, then, should the quality of an assessment be judged? And which assessments should be subjected to such an analysis?

While the criticisms have been directed most frequently at the perceived weaknesses and inadequacies of EPA assessments, those of other federal agencies have received attention as well. However, we would argue that the standards of performance demanded of EPA and other federal agencies also should be demanded of state and local government agencies, communities, regulated industry, public interest groups, academics and any other parties which conduct or fund risk assessments and related research on their own behalf. Their products also should be subjected to external expert peer review and public comment and a quality analysis. In other words, it should not only be EPA and other federal agencies that are obligated to upgrade their assessment processes and practices and prove their credibility. It is incumbent upon government's involved stakeholders to do the same, especially the regulated community. In particular, if a non-governmental entity wishes to have its assessments be considered in the governmental regulatory decision-making process, then its products should be judged in the same rigorous manner expected of government.

Transparency, effectiveness, efficiency and scientific integrity are all essential traits which are captured in the Guide. These features are applicable to any type of assessment, whether it encompasses just one, or two, or all four phases of the risk assessment paradigm, whether qualitative or quantitative, screening level or highly sophisticated and complex. These characteristics apply to both traditional approaches as well as to the newer 21st century or "next generation" approaches, as described, for example, in NRC 2007, 2012 and US EPA 2014b.

Organizations, government and otherwise, rightfully should be held accountable to their respective constituencies. A carefully-crafted set of performance measures can serve as a credible tool for determining the level of success in meeting performance expectations. In the

case of the U.S. federal government, each agency of the Executive branch is required to submit annual performance reports to the Office of Management and Budget (OMB) (OMB 2014). These reports represent data-driven reviews of the strategic objectives established in their respective Strategic Plans and include an articulation of achievements made in meeting program objectives along with identification of areas where improvement may be needed. Among the purposes they serve are: informing long-term strategic decision-making; facilitating identification and adoption of opportunities for improvement; identifying areas where additional evaluation, other studies or analyses of data are needed; identifying where additional skills or other capacity are needed; strengthening collaboration on crosscutting issues, and improving transparency. Similar activities carried out in the private and non-profit sectors are seen to be reflective of good management practices. Progress is measured against a set of pre-determined performance criteria. We see the value of employing the Guide as an application of the same concepts to an evaluation procedure for judging the quality of assessments. Given the growing consensus among the many parties of what characteristics a high-value assessment should possess, it is useful to have a guide, available as a single document, to provide direction for authors, decision-makers, reviewers, readers and other users when they are judging the quality of such an assessment.

What key elements might one document include in order to determine whether or not an assessment meets the criteria for a high quality product? The Guide contains a series of points focused on good scientific practices, as gleaned from the expert panel reports to be used in developing credible and transparent assessments. We acknowledge that the measures presented

may not be all inclusive, but believe that they capture the key considerations for a high quality assessment. We highlight a few overarching themes around the points captured in the Guide.

### **Designing with Focus**

The Guide begins with the foremost characteristic of a high quality assessment: fit-for-purpose. That is, it clearly addresses the problem(s) and questions at hand and considers the options or boundaries for which decisions need to be made. Before an assessment is initiated, problem formulation, planning and scoping must occur. This is a crucial step for an effective and efficient assessment. A number of points must be addressed, such as the overall purpose and general scope of the assessment, the assessment products needed to inform decision making, the resources required, who the authors of the risk assessment will be and their respective roles, the time table, *etc.* (US EPA 2014a).

Good problem definition needs to address the issues and concerns of the key participants and stakeholders. Critical to public confidence and a successful product is an open process that allows early and continuing dialogue with the stakeholder community. Stakeholders can serve a valuable role in identifying issues, data, and alternative approaches to conducting an assessment.

It is worthwhile noting that organizations usually are faced with finite resources and time to conduct their assessments; thus, not only are there scientific drivers that demand improved quality, but also the realization that resources must be used efficiently. The extent of documentation needs to be balanced by resources and priorities, especially when the timeliness of the response is critical (Health Canada 2000; Dellarco et al. 2010). The mere presence of a substance in the environment does not necessarily mean that it poses a threat to human health or

the environment; thus, an approach that considers exposure early on can better focus resources on those stressors which pose exposure scenarios of concern. Some NRC reports (NRC 1996, 2009) and other publications (Pastoor et al. 2014), also have noted that problem formulation must include an early consideration of the relevant exposure scenarios/pathways along with potential options for managing or mitigating the exposures. Only then will the assessment efficiently and effectively serve the needs of the user. In all cases, transparency is key; the selected approaches should be well-described in the problem formulation document.

### **Selecting Data and Evaluating Reliability**

Once the problem formulation phase is completed, a systematic review process should follow. An important aspect of this process is the definition and documentation of the search and review procedures employed to ensure transparency and that the results can be replicated. The literature search and procedure for data collection and evaluation will shape the scientific basis of the assessment, and thus needs to be guided by the questions, goals, and methodologies identified in the problem formulation phase. Keep in mind that this is an iterative process. The systematic review process also is designed to reveal and minimize bias. Although bias is difficult to eliminate completely, if all documents (including the documented search for studies) have been made publically available, and the reasoning behind choices made clear, then bias at least can be identified, addressed, and minimized. Also, having an assessment team of multi-disciplinary scientists can ensure a range of perspectives that permits alternative interpretations to be considered along with the evidence that supports or refutes these alternatives.

Only well-documented epidemiologic, toxicity and exposure studies based on reliable data should be used-- particularly in an assessment that could have a significant impact. Studies that

are poorly documented or those with questionable study design and repeatability should be identified as such and not used. These judgments would become clear in the course of conducting a scientifically rigorous and transparent systematic review. The evaluation of a study's quality is essential to the weight-of-evidence process. In 2002, OMB issued guidelines that "provide policy and procedural guidance to federal agencies for ensuring and maximizing the quality, objectivity, utility, and integrity of information (OMB 2002). Each federal agency was required to issue implementing guidelines (*e.g.*, HHS 2002; US EPA 2002). Foreign governments and international organizations also have developed their own quality guidelines (OECD 2007, 2012; ECHA 2011; Health Canada 2015).

### **Combining Evidence**

Once the candidate studies have been assembled and evaluated for quality, the relevant credible body of data is subjected to a weight-of-evidence (WoE)/evidence integration exercise in order to characterize the extent to which the hypotheses put forward are or are not supported. This aspect of risk assessment has been the most difficult to execute well. While authors may assert that they have weighed and integrated all of the information in a constructive manner, their interpretive approach and how it was applied often is unclear or lacking in documentation (NAS 2011). If the integration process is based upon well-defined criteria that ensure structure and rigor, it can be of important value and scientific use. Rhomberg et al. (2013) reviewed nearly 50 frameworks to evaluate best practices for WoE analyses for determining causation of chemical risks, concluding that this review, along with its companion workshop deliberations, provides "actionable best practices recommendations that can be put to immediate use in designing and conducting systematic reviews/WoE determinations for chemical hazards/risks."

Separately, Rhomberg and his colleagues have developed a hypothesis-based approach for synthesizing dissimilar and complex data sets to more successfully support a true WoE output (Rhomberg, et al. 2010; Rhomberg, 2014/2015). Other hypothesis-based WoE approaches (using Bradford Hill-like considerations) have been developed which also promote a systematic evaluation and integration of data, in this case to characterize toxic modes of action and their relevance to human targets. The mode-of-action/human relevance framework developed by WHO/IPCS and the International Life Sciences Institute focuses on human health (Sonich-Mullin et al. 2001; Meek et al. 2003; Seed et al. 2005; Boobis et al. 2006, 2008). Many examples exist of this framework's application and incorporation of its results into the decision-making process by regulatory authorities, both domestic (US EPA 2005, 2010 ; CalEPA 2013) and international (FAO/WHO 2012, 2013; EFSA 2014a, b). The WHO/IPCS framework has evolved to incorporate the use of quantitative dose-response and temporal relationships for key events within a mode-of-action to reduce uncertainty and better inform quantitative risk assessment (Julien et al. 2009; Simon et al. 2014) along with other refinements as experience in its application accrues (Meek et al. 2014a, b). In its development of a library of Adverse Outcome Pathways (AOPs), the OECD is coordinating its activities with WHO/IPCS, and has incorporated the weight of evidence approach of the WHO/IPCS framework into its guidance and template to assess the evidence in support of an AOP (OECD 2013). Conceptually similar to mode of action, an AOP is defined as "an analytical construct that describes a sequential chain of causally linked events at different levels of biological organisation that lead to an adverse [human] health or ecotoxicological effect. AOPs are the central element of a toxicological knowledge framework being built to support chemical risk assessment based on mechanistic reasoning" (OECD 2013). At present, AOPs are being developed to address the common goal

of identifying faster, more reliable interpretable methods (*e.g.*, *in vitro* screens, strengthening read-across methods and QSARs). Understanding of AOPs, in turn, informs mode of action analyses for specific chemicals or groups of chemicals.

### **Ensuring Transparency and Clarity**

It is essential that an assessment be transparent (*i.e.*, possess openness in procedural process and scientific aspects) but also easily understood (clarity). Scientific assessments can be hard to follow, even for a technical audience. However, it should never be assumed that difficulties in understanding the assessment are simply a result of the complexity of the scientific data used, the type of analysis, and the concepts applied. Rather, they also can be the result of deficiencies in communication. Clarity is an important feature of any quality assessment. Mere presentation of information is insufficient for an assessment intended to support decision making. The points at which choices are made in the selection of data, use of defaults and assumptions, consideration of alternative methods, *etc.*, and the reasoning underlying these choices need to be clearly captured for users and other readers. The inclusion of an executive summary written for both technical and nontechnical audiences, and summary tables and figures can facilitate this. Clarity in the assessment is particularly important in situations involving difficult determinations about data interpretation or conflicting views on plausible alternative methods.

Uncertainty is a scientific reality that cannot be totally eliminated, and, therefore, must be acknowledged and explained, including its impact on the risk conclusions and estimates. It is essential that all conclusions and risk estimates (including alternatives) be described explicitly in the context of certainties and uncertainties. Uncertainties that are identified should be based on empirical knowledge rather than speculative “unknown negative effects.” In general, the

nature of any uncertainty raised should be addressed, at least qualitatively, if not quantitatively. It is important for an assessment to point out ways that uncertainty could be reduced. Progress has been made in the tools to characterize uncertainty (particularly for exposure) and guidance is available (*e.g.*, WHO IPCS 1999, 2009; US EPA 2000). Nonetheless, there remains a need to strengthen this area of assessment, as noted in the NRC report *Environmental Decisions in the Face of Uncertainty* (NRC 2013).

### Objectivity and Reasonableness

There are a number of potential sources of bias that can occur in the various phases of the assessment process (*e.g.*, study design, data selection, data interpretation, choice of defaults, models, methods). One area of bias concerns author bias (*i.e.*, of those conducting the assessment). It should be acknowledged here that scientists/risk assessors have biases, beliefs, and opinions; however, it is critical that advocacy positions not infiltrate and influence an assessment. A credible assessment that embodies best scientific practices is based upon empirical evidence and ensures scientific objectivity.

While remaining mindful of the goal to protect human health and the environment at reasonable cost, a truly useful assessment should reflect common-sense application of assumptions and policies and avoid mischaracterizations of hazard or unrealistic estimations of exposure and risk (in either direction). A reality check is important to ensure that conservative assumptions have not over-accumulated in the assessment. Scientifically-sound characterization of hazard and more realistic estimates of exposure and risk lend credibility to, and improve confidence in, the assessment. Related to this point is the consideration of alternative conclusions or risk estimates that have reasonable evidence for support. Presenting supportable alternatives

provides more information for consideration in decision-making and lends more credibility to the assessment. This is an important responsibility of the authors of an assessment. The regulatory process demands consideration of the full range of possibilities. Robust characterization of all of the supportable alternatives is critical to the users of the assessments, especially decision-makers, and is an essential characteristic of a credible assessment.

### Peer Reviewing the Science

Scientific peer participation and peer review are key elements of the assessment process. They play a critical role in ensuring the credibility and integrity of the scientific information generated, evaluated, and communicated by the authors. Peer participation includes involvement in the development of an assessment. Independent peer review of a draft assessment can be a reliable judge of the usefulness, quality and relevance of the assessment; it also can evaluate the scientific objectivity and the consideration of alternative interpretations and methods.

Agencies of the federal government are expected to develop and execute a formal peer review process. This is mandated, in part, by OMB. Its Peer Review Bulletin established “government-wide guidance aimed at enhancing the practice of peer review of government science documents” (OMB 2004). The Bulletin also required that each federal agency develop and publish a process for conducting peer reviews that is transparent, incorporating the Bulletin-established minimum standards for when and what type of peer review is required and appropriate for a given circumstance. In addition, if a peer review is being conducted as a component of activities covered by the Federal Advisory Committee Act (FACA), additional steps may need to be incorporated into the process (US House of Representatives 1972). FACA

provisions also apply to committees convened by the National Academies of Sciences (NAS 2003).

A number of state governments also have developed policies and guidance with regard to peer review. Some practices are mandated under state law (e.g., Rowe 2006). Some state governments convene standing advisory panels/committees which may conduct peer review of their government's assessments or develop their own assessments. These panels generally convene in a public forum and/or provide opportunities for interested stakeholders to provide written review and comment before finalizing their recommendations to their respective state. It is less clear as to what peer review policies and practices may exist in local governments. Nonetheless, there is great demand that all levels of government function in an open and transparent manner and provide for stakeholder input at various points in the decision-making process.

On the other hand, there is no obligation to perform the same kind of open and collaborative peer review on the part of the regulated community, public interest groups, academics or other parties which prepare or fund assessments or related research on their own behalf. Rarely, their assessments are peer-reviewed by a panel convened by a third party. However, these panels generally do not perform their evaluations in a public setting, do not solicit public comment, and their membership and deliberations may or may not be made public after the fact.

Authors/sponsors may opt for publication of their assessment or research in a scientific peer-reviewed journal. This serves two purposes: 1) the information is peer-reviewed, presumably by qualified experts, and 2) the final product can be widely-disseminated to the interested audience. However, the journal peer review process is generally not very transparent. While a

reader may have access to the final publication, who performed the peer review or how the authors responded to any comments received is seldom revealed. In many cases, there is not access to sufficient information for the reader to attempt a replication of the assessment or research study, although some journals are now providing the opportunity for authors to provide supplementary information with their manuscript. Further improvements in the journal publishing arena may arise in light of the issuance of the Principles and Guidelines for Reporting Preclinical Research, agreed upon in a gathering convened by NIH, *Nature* and *Science* of more than 30 major journal editors, representatives from funding agencies and scientific leaders (NIH 2015a). Application of these Principles and Guidelines will serve to guide and enhance the development of a harmonized, systematic review process. Of particular value are the recommendations made in the areas of rigorous statistical analysis, transparency in reporting, and data and material sharing. The effort with regard to data sharing already has begun within the federal government at the direction of the Office of Science Technology and Policy. A memo issued in February 2013 directs “each Federal agency with over \$100 million in annual conduct of research and development expenditures to develop a plan to support increased public access to the results of research funded by the Federal Government. This includes any results published in peer-reviewed scholarly publications that are based on research that directly arises from Federal funds” (OSTP 2013). To date, at least 14 agencies have finalized their plans (*e.g.*, US FDA 2015; NIH 2015b). EPA’s final public access plan has yet to be issued.

### **Applying the Guide**

The Guide is intended to be both a self-assessment tool and a tool to enable others to judge the robustness of an assessment. We would expect that any individual employing the Guide to judge the quality of an assessment would do so in accordance with the policies and practices of

the authoring organization, informed by his/her experience, knowledge and perspectives, while remaining vigilant to minimize any personal bias. There may well be situations where an individual possesses or is perceived to possess a conflict of interest, and, thus, may not be the best candidate to perform the quality review. The Guide also could serve as an especially valuable companion piece to the charge that peer reviewers receive from a sponsoring organization when they are asked to evaluate an assessment. Although an assessment may not “tick every box” for all elements in the Guide, this does not necessarily mean that the assessment is not high quality. Whether or not an assessment possesses all or just some of the attributes identified in the Guide will depend upon the nature of the assessment (*e.g.*, screening vs. in-depth or encompassing just one or more elements of the risk assessment paradigm). This is determined by the scope as developed during problem formulation. In any case, while some attributes may not be present, having this noted in the application of the Guide allows for enhanced transparency. If the Guide application results are released along with the assessment, readers can easily see which attributes were incorporated and which were not. This level of transparency will be useful to decision-makers, peer reviewers, and other interested parties.

## **Conclusions**

A review of 30+ years of NAS/NRC/IOM and other federal government-funded expert panel reports coupled with government agency science and policy guidance developed in response to these reports reveals the degree of consensus around the key elements that a high quality and useful assessment should possess. These key elements have been used to develop a Guide by which decision-makers can determine if they have the highest quality information available to

carry out their responsibilities and authors, sponsors, risk assessors, peer reviewers and other stakeholders can determine if an assessment “measures up” to current best scientific practices.

Over the years, many reports have been published concerning the quality, transparency and usefulness for decision-making of human health and ecological assessments prepared by agencies of the U.S. federal government. Recommendations for improvement have been offered on every aspect of the design, content, execution and role of risk assessments in the decision-making process. From the plethora of observations and recommendations, a consensus is evolving as to what characteristics an assessment should possess to be deemed of high quality. Each federal agency is mandated to conduct data-driven performance reviews of their priorities to gauge their progress towards achieving their stated goals (OMB 2014).

Successfully-functioning state and local governments and organizations in the private, non-profit and academic sectors also often incorporate such reviews as a matter of good management practice. Capitalizing on this precedent, we have created a Guide that provides the author, decision-maker, risk assessor, peer reviewer or other interested stakeholder a means for determining whether or not a particular assessment meets the criteria for excellence. We envision it to be especially useful as a self-evaluation tool and a companion piece to consult during the problem formulation process. It can also serve as a complementary component of the charge that peer reviewers receive from a sponsoring organization when they are asked to evaluate an assessment. We believe it would be prudent to use the Guide (or a modified version which is tailored appropriately for its scope and for a technical peer review) as a gauge for determining the quality of an assessment, in advance of circulating the assessment to a larger audience for review and comment, and before any regulatory decisions are made.

The Guide, constituting a series of statements/measures, covers many aspects of problem-formulation, systematic review of the literature, hazard identification/characterization, dose-response assessment/characterization, risk characterization and peer review. The statements capture key aspects of the essential traits of an assessment around which consensus has emerged. When the totality of applicability to the statements is viewed holistically, as in a weight-of-evidence evaluation, one should be able to easily gauge the level of accomplishment in meeting the objectives of creating a good assessment. Simply put, the more often a characteristic of the Guide can be ascribed to the assessment being judged, the more likely it will be seen as meeting its performance expectations.

## References

- Boobis AR, Cohen SM, Dellarco V, McGregor D, Meek ME, Vickers C, Willcocks D, Farland W. 2006. IPCS framework for analyzing the relevance of a cancer mode of action for humans. *Crit. Rev. Toxicol.* 36:781-792.
- Boobis AR, Doe JE, Heinrich-Hirsch B, Meek ME, Munn S, Ruchirawat M, Schlatter J, Seed J, Vickers C. 2008. IPCS framework for analyzing the relevance of a noncancer mode of action for humans. *Crit. Rev. Toxicol.* 38:87-96.
- CalEPA (California Environmental Protection Agency). 2013. Reproductive and Cancer Hazard Assessment Branch. Office of Environmental Health Hazard Assessment. Evidence on the carcinogenicity of diisononyl phthalate (DINP). Available: [http://oehha.ca.gov/Prop65/hazard\\_ident/pdf\\_zip/DINP\\_HID100413.pdf](http://oehha.ca.gov/Prop65/hazard_ident/pdf_zip/DINP_HID100413.pdf). [accessed 17 January 2016].
- Dellarco V, Henry T, Sayre P, Seed J, Bradbury S. 2010. Meeting the common needs of a more effective and efficient testing and assessment paradigm for chemical risk management. *J Toxicol Environ Health B Crit Rev.* 13(2-4):347-60.
- ECHA (European Chemicals Agency). 2011. Guidance on information requirements and chemical safety assessment. Chapter R.4: Evaluation of available information. Version 1.1. Available: [https://echa.europa.eu/documents/10162/13643/information\\_requirements\\_r4\\_en.pdf](https://echa.europa.eu/documents/10162/13643/information_requirements_r4_en.pdf) [accessed 17 January 2016].
- EFSA (European Food Safety Agency). 2010. Application of systematic review methodology to food and feed safety assessments to support decision making. *EFSA Journal* 8(6):1637 [90 pp.]. Available: <http://www.efsa.europa.eu/en/efsajournal/pub/1637>. [accessed 17 January 2016].
- EFSA (European Food Safety Agency). 2011. Submission of scientific peer-reviewed open literature for the approval of pesticide active substances under Regulation (EC) No 1107/2009. *EFSA Journal* 9(2):2092 [49 pp.]. Available: <http://www.efsa.europa.eu/en/efsajournal/pub/2092>. [accessed 17 January 2016].

- EFSA (European Food Safety Agency). 2013. International Frameworks Dealing with Human Risk Assessment of Combined Exposure to Multiple Chemicals. EFSA Journal 11(7):3313 [69 pp]. Available: [http://www.efsa.europa.eu/sites/default/files/scientific\\_output/files/main\\_documents/3313.pdf](http://www.efsa.europa.eu/sites/default/files/scientific_output/files/main_documents/3313.pdf). [accessed 17 January 2016].
- EFSA (European Food Safety Authority). 2014a. Reasoned opinion on the review of the existing maximum residue levels (MRLs) for clothianidin and thiamethoxam according to Article 12 of Regulation (EC) No 396/20051. EFSA Journal 12(12):3918. Available: [http://www.efsa.europa.eu/sites/default/files/scientific\\_output/files/main\\_documents/3918.pdf](http://www.efsa.europa.eu/sites/default/files/scientific_output/files/main_documents/3918.pdf). [accessed 17 January 2016].
- EFSA (European Food Safety Authority). 2014b. Reasoned opinion on the review of the existing maximum residue levels (MRLs) for thiophanate-methyl and carbendazim according to Article 12 of Regulation (EC) No 396/20051. EFSA Journal 12(12):3919. Available: [http://www.efsa.europa.eu/sites/default/files/scientific\\_output/files/main\\_documents/3919.pdf](http://www.efsa.europa.eu/sites/default/files/scientific_output/files/main_documents/3919.pdf). [accessed 17 January 2016].
- FAO/WHO (Food and Agriculture Organization/World Health Organization). 2012. 2011 Joint Meeting of the FAO Panel of Experts on Pesticide Residues in Food and the Environment and the WHO Core Assessment Group on Pesticide Residues. Pesticide residues in food - 2011. Part II-Toxicological evaluations. Etofenprox. pp. 253-324. WHO Press. Geneva, Switzerland. Available: [http://apps.who.int/iris/bitstream/10665/75147/1/9789241665278\\_eng.pdf](http://apps.who.int/iris/bitstream/10665/75147/1/9789241665278_eng.pdf). [accessed 17 January 2016].
- FAO/WHO (Food and Agriculture Organization/World Health Organization). 2013. 2012 Joint Meeting of the FAO Panel of Experts on Pesticide Residues in Food and the Environment and the WHO Core Assessment Group on Pesticide Residues. Pesticide residues in food - 2012. Part II-Toxicological evaluations. Fluxapyroxad. pp. 363-545. WHO Press. Geneva, Switzerland. Available: <http://www.who.int/foodsafety/publications/jmpr-monographs/en/>. [accessed 17 January 2016].

- Health Canada. 2000. Health Canada Decision-Making Framework for Identifying, Assessing, and Managing Health Risks. Ottawa, Canada. Available: [http://www.hc-sc.gc.ca/ahc-asc/pubs/hpfb-dgpsa/risk-risques\\_tc-tm-eng.php](http://www.hc-sc.gc.ca/ahc-asc/pubs/hpfb-dgpsa/risk-risques_tc-tm-eng.php). [accessed 17 January 2016].
- Health Canada. 2015. Health Canada Regulatory Transparency and Openness Framework and Action Plan 2015-2018. Ottawa, Canada. Available: <http://www.hc-sc.gc.ca/home-accueil/rto-tor/2015-18-fap-cpa-eng.php>. [accessed 17 January 2016].
- HHS (Department of Health and Human Services). 2002. HHS Guidelines for Ensuring and Maximizing the Quality, Objectivity, Utility, and Integrity of Information Disseminated to the Public. Washington, DC. Available: <http://aspe.hhs.gov/infoquality/Guidelines/index.shtml>. [accessed 17 January 2016].
- Higgins JPT, Green S (editors). 2011. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]. The Cochrane Collaboration. Available: [www.cochrane-handbook.org](http://www.cochrane-handbook.org). [accessed 17 January 2016].
- IOM (Institute of Medicine). 2011. Finding What Works in Health Care: Standards for Systematic Reviews. National Academies Press, Washington, DC. Available: [http://www.nap.edu/download.php?record\\_id=13059](http://www.nap.edu/download.php?record_id=13059). [accessed 17 January 2016].
- Julien E, Boobis AR, Olin SS, 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: 682-689.
- Meek ME, Boobis A, Cote I, Dellarco V, Fotakis G, Munn S, Seed J, Vickers C. 2014a. New developments in the evolution and application of the WHO/IPCS framework on mode of action/species concordance analysis. *J. Appl. Toxicol.* 34:1-18
- Meek ME, Bucher JR, Cohen SM, Dellarco V, Hill RN, Lehman-McKeeman LD, Longfellow DG, Pastoor T, Seed J, Patton DE. 2003. A framework for human relevance analysis of information on carcinogenic modes of action. *Crit. Rev. Toxicol.* 33:591-653.
- Meek ME, Palermo CM, Bachman AN, North CM, Lewis RJ. 2014b. Mode of Action Human Relevance (MOA/HR) Framework – Evolution of the Bradford Hill Considerations and Comparative Analysis of Weight of Evidence. *J. Appl. Toxicol.* 34(6):595-606

- NAS (National Academies of Sciences). 2003. Policy and Procedures on Committee Composition and Balance and Conflicts of Interest for Committees Used in the Development of Reports. Available: <http://www.nationalacademies.org/coi/>. [accessed 17 January 2016].
- NIH (National Institutes of Health). 2015a. Principles and Guidelines for Reporting Preclinical Research. Bethesda, MD. Available: <http://www.nih.gov/research-training/rigor-reproducibility/principles-guidelines-reporting-preclinical-research>. [accessed 17 January 2016].
- NIH (National Institutes of Health). 2015b. National Institutes of Health Plan for Increasing Access to Scientific Publications and Digital Scientific Data from NIH Funded Scientific Research. Bethesda, MD. Available: <http://grants.nih.gov/grants/NIH-Public-Access-Plan.pdf>. [accessed 17 January 2016].
- NRC (National Research Council). 1983. Risk Assessment in the Federal Government: Managing the Process. National Academies Press, Washington, DC. Available: <http://www.nap.edu/catalog/366/risk-assessment-in-the-federal-government-managing-the-process>. [accessed 17 January 2016].
- NRC (National Research Council). 1994. Science and Judgment in Risk Assessment. National Academies Press, Washington, DC. Available: <http://www.nap.edu/catalog/2125/science-and-judgment-in-risk-assessment>. [accessed 17 January 2016].
- NRC (National Research Council). 1996. Understanding Risk: Informing Decisions in a Democratic Society. National Academies Press, Washington, DC. Available: <http://www.nap.edu/catalog/5138/understanding-risk-informing-decisions-in-a-democratic-society>. [accessed 17 January 2016].
- NRC (National Research Council). 2007. Toxicity Testing in the 21st Century: A Vision and a Strategy. National Academies Press, Washington, DC. Available: <http://www.nap.edu/catalog/11970/toxicity-testing-in-the-21st-century-a-vision-and-a>. [accessed 17 January 2016].

- NRC (National Research Council). 2009. Science and Decisions: Advancing Risk Assessment. National Academies Press, Washington, DC. Available: <http://www.nap.edu/catalog/12209/science-and-decisions-advancing-risk-assessment>. [accessed 17 January 2016].
- NRC (National Research Council). 2011. Review of the Environmental Protection Agency's Draft IRIS Assessment of Formaldehyde. National Academies Press, Washington, DC. Available: <http://www.nap.edu/catalog/13142/review-of-the-environmental-protection-agencys-draft-iris-assessment-of-formaldehyde>. [accessed 17 January 2016].
- NRC (National Research Council). 2012. Exposure Science in the 21st Century: A Vision and a Strategy. National Academies Press, Washington, DC. Available: <http://www.nap.edu/catalog/13507/exposure-science-in-the-21st-century-a-vision-and-a>. [accessed 17 January 2016].
- NRC (National Research Council). 2013. Environmental Decisions in the Face of Uncertainty. National Academies Press, Washington, DC. Available: <http://www.nap.edu/catalog/12568/environmental-decisions-in-the-face-of-uncertainty>. [accessed 17 January 2016].
- NRC (National Research Council). 2014. Review of EPA's Integrated Risk Information System (IRIS) Process. National Academies Press, Washington, DC. Available: <http://www.nap.edu/catalog/18764/review-of-epas-integrated-risk-information-system-iris-process>. [accessed 17 January 2016].
- NTP (National Toxicology Program). 2015. Office of Health Assessment and Translation. Handbook for Conducting a Literature-Based Health Assessment Using OHAT Approach for Systematic Review and Evidence Integration. National Institute of Environmental Health Sciences, Research Triangle Park, NC. Available: [http://ntp.niehs.nih.gov/ntp/ohat/pubs/handbookjan2015\\_508.pdf](http://ntp.niehs.nih.gov/ntp/ohat/pubs/handbookjan2015_508.pdf). [accessed 17 January 2016].
- OECD (Organisation of Economic Co-operation and Development). 2007. OECD Principles and Guidelines for Access to Research Data from Public Funding. Paris, France. Available: <http://www.oecd.org/sti/sci-tech/38500813.pdf>. [accessed 17 January 2016].

- OECD (Organisation of Economic Co-operation and Development). 2012. Quality Framework and Guidelines for OECD Statistical Activities. Version 2011/1. Paris, France. Available: [www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=std/qfs\(2011\)1&doclanguage=en](http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=std/qfs(2011)1&doclanguage=en). [accessed 17 January 2016].
- OECD (Organisation of Economic Co-operation and Development). 2013. Guidance Document on Developing and Assessing Adverse Outcome Pathways. Series on Testing and Assessment No. 184. ENV/JM/MONO(2013)6. Paris, France. Available: [www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=env/jm/mono\(2013\)6&doclanguage=en](http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=env/jm/mono(2013)6&doclanguage=en). [accessed 17 January 2016].
- OMB (Office of Management and Budget). 2002. Guidelines for Ensuring and Maximizing the Quality, Objectivity, Utility, and Integrity of Information Disseminated by Federal Agencies. Executive Office of the President. Available: [https://www.whitehouse.gov/omb/fedreg\\_final\\_information\\_quality\\_guidelines/](https://www.whitehouse.gov/omb/fedreg_final_information_quality_guidelines/). [accessed 17 January 2016].
- OMB (Office of Management and Budget). 2004. Final Information Quality Bulletin for Peer Review. Executive Office of the President. Washington, DC. Available: [https://www.whitehouse.gov/omb/memoranda\\_fy2005\\_m05-03/](https://www.whitehouse.gov/omb/memoranda_fy2005_m05-03/). [accessed 17 January 2016].
- OMB (Office of Management and Budget). 2014. Circular A-11, Part 6. Preparation and Submission of Strategic Plans, Annual Performance Plans, and Annual Performance Reports. Executive Office of the President. Washington, DC. Available: <https://www.whitehouse.gov/sites/default/files/omb/performance/omb-circular-a11-part-6.pdf>. [accessed 17 January 2016].
- OSTP (Office of Science Technology and Policy). 2013. Increasing Access to the Results of Federally Funded Scientific Research. Executive Office of the President. Washington, DC. Available: [https://www.whitehouse.gov/sites/default/files/microsites/ostp/ostp\\_public\\_access\\_memo\\_2013.pdf](https://www.whitehouse.gov/sites/default/files/microsites/ostp/ostp_public_access_memo_2013.pdf). [accessed 17 January 2016].

- Pastoor TP, Bachman AN, Bell DR, Cohen SM, Dellarco M, Dewhurst IC, Doe JE, Doerrer NG, Embry MR, Hines RN, Moretto A, Phillips RD, Rowlands JC, Tanir JY, Wolf DC, Boobis AR. 2014. A 21st century roadmap for human health risk assessment. *Crit Rev Toxicol*, 44(S3):1–5.
- Presidential/Congressional Commission on Risk Assessment and Risk Management. 1997. Framework for Environmental Health Risk Management. Washington, DC. Available: <http://cfpub.epa.gov/ncea/cfm/recordisplay.cfm?deid=55006>. [accessed 17 January 2016].
- Rhomberg LR. 2014/2015. Hypothesis-Based Weight of Evidence: An Approach to Assessing Causation and its Application to Regulatory Toxicology. *Risk Anal*. doi:10.1111/risa.12206 and [Risk Anal](#). 2015 35(6):1114-24.
- Rhomberg LR, Bailey LA, Goodman JE. 2010. Hypothesis-based weight of evidence: A tool for evaluating and communicating uncertainties and inconsistencies in the large body of evidence in proposing a carcinogenic mode of action—naphthalene as an example. *Crit. Rev. Toxicol*. 40(8): 671–696.
- Rhomberg LR, Goodman JE, Bailey LA, Prueitt RL, Beck NB, Bevan C, Honeycutt M, Kaminski NE, Paoli G, Pottenger LH, Scherer RW, Wise KC, Becker RA. 2013. A survey of frameworks for best practices in weight-of-evidence analyses. *Crit Rev Toxicol*. 43(9):753-84.
- Rowe, GW. 2006. California Environmental Protection Agency. External Scientific Peer Review Guidelines. Available: [http://www.swrcb.ca.gov/water\\_issues/programs/peer\\_review/docs/exhibit\\_f.pdf](http://www.swrcb.ca.gov/water_issues/programs/peer_review/docs/exhibit_f.pdf). [accessed 17 January 2016].
- Seed J, Carney EW, Corley RA, Crofton KM, DeSesso JM, Foster PM, Kavlock R, Kimmel G, Klaunig J, Meek ME, Preston RJ, Slikker WJ, Tabacova S, Williams GM, Wiltse J, Zoeller RT, Fenner-Crisp P, Patton DE. 2005. Overview: Using mode of action and life stage information to evaluate the human relevance of animal toxicity data. *Crit. Rev. Toxicol*. 35:664-672.

- Simon T, Simons, Jr, SS, Preston RJ, Boobis AR, Cohen SM, Doerrner NG, Fenner-Crisp PA, McMullin T, McQueen C, Rowlands JC, HESI RISK21 Dose-Response Subteam. 2014. The use of mode of action information in risk assessment: quantitative key events/dose-response framework for modeling the dose-response for key events. *Crit Rev Toxicol.* 44 Suppl 3:17-43.
- Sonich-Mullin C, Fielder R, Wiltse J, Baetcke K, Dempsey J, Fenner-Crisp P, Grant D, Hartley M, Knaap A, Kroese D, Mangelsdorf I, Meek E, Rice JM, Younes M, International Programme on Chemical Safety. 2001. IPCS conceptual framework for evaluating a mode of action for chemical carcinogenesis. *Regul. Toxicol. Pharmacol.* 34:146-152.
- TCEQ (Texas Commission on Environmental Quality). 2014. TCEQ Recommendations for Systematic Review and Evidence Integration. Proposed Position Paper. Austin, TX. Available: [http://www.tceq.com/assets/public/implementation/tox/dsd/proposed/nov14/systematic\\_review.pdf](http://www.tceq.com/assets/public/implementation/tox/dsd/proposed/nov14/systematic_review.pdf). [accessed 17 January 2016].
- US EPA (U.S. Environmental Protection Agency). 1992. Framework for Ecological Risk Assessment. Risk Assessment Forum. Washington, DC. EPA/630/R-92/001. Available: <http://www2.epa.gov/osa/framework-ecological-risk-assessment>. [accessed 17 January 2016].
- US EPA (U.S. Environmental Protection Agency). 1993. William D. Ruckelshaus: Oral History Interview. EPA 202-K-92-0003. January 1993. Interview conducted by Dr. Michael Gorn. Available: <http://www2.epa.gov/aboutepa/william-d-ruckelshaus-oral-history-interview>. [accessed 17 January 2016].
- US EPA (U.S. Environmental Protection Agency). 1998. Guidelines for Ecological Risk Assessment. Risk Assessment Forum. EPA/630/R-95/002F. Washington, DC. Available: [http://www2.epa.gov/sites/production/files/2014-11/documents/eco\\_risk\\_assessment1998.pdf](http://www2.epa.gov/sites/production/files/2014-11/documents/eco_risk_assessment1998.pdf). [accessed 17 January 2016].

- US EPA (U.S. Environmental Protection Agency). 2000. Science Policy Council handbook: Risk characterization. Office of Science Policy, Office of Research and Development, Washington, DC; EPA 100-B-00-002. Available: [http://www.epa.gov/sites/production/files/2015-10/documents/osp\\_risk\\_characterization\\_handbook\\_2000.pdf](http://www.epa.gov/sites/production/files/2015-10/documents/osp_risk_characterization_handbook_2000.pdf). [accessed 17 January 2016].
- US EPA (U.S. Environmental Protection Agency). 2002. Guidelines for Ensuring and Maximizing the Quality, Objectivity, Utility, and Integrity of Information Disseminated by the Environmental Protection Agency. Office of Environmental Information. Washington, DC. EPA/260R-02-008. Available: <http://www.epa.gov/sites/production/files/2015-08/documents/epa-info-quality-guidelines.pdf>. [accessed 17 January 2016].
- US EPA (U.S. Environmental Protection Agency). 2003. Framework for Cumulative Risk Assessment. Risk Assessment Forum. U.S. Environmental Protection Agency. Washington, DC. EPA/630/P-02/001F. Available: [http://www.epa.gov/sites/production/files/2014-11/documents/frmwrk\\_cum\\_risk\\_assmnt.pdf](http://www.epa.gov/sites/production/files/2014-11/documents/frmwrk_cum_risk_assmnt.pdf). [accessed 17 January 2016].
- US EPA (U.S. Environmental Protection Agency). 2005. Science Issue Paper: Mode of Carcinogenic Action for Cacodylic Acid (Dimethylarsinic Acid, DMAV) and Recommendations for Dose Response Extrapolation. Health Effects Division, Office of Pesticide Programs, US EPA, Washington, DC. Available: [http://archive.epa.gov/pesticides/reregistration/web/pdf/dma\\_moa-2.pdf](http://archive.epa.gov/pesticides/reregistration/web/pdf/dma_moa-2.pdf). [accessed 17 January 2016].
- US EPA (U.S. Environmental Protection Agency). 2006. 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. Available: <http://cfpub.epa.gov/ncea/risk/recordisplay.cfm?deid=158363>. [accessed 17 January 2016].

- US EPA (U.S. Environmental Protection Agency). 2010. Toxicological Review of Acrylamide (CAS No. 79-06-1) In Support of Summary Information on the Integrated Risk Information System (IRIS). EPA/635/R-07/009F. Washington, DC. Available: [http://cfpub.epa.gov/ncea/iris/iris\\_documents/documents/toxreviews/0286tr.pdf](http://cfpub.epa.gov/ncea/iris/iris_documents/documents/toxreviews/0286tr.pdf). [accessed 17 January 2016].
- US EPA (U.S. Environmental Protection Agency). 2012. Guidance for Identifying, Selecting and Evaluating Open Literature Studies. Office of Pesticide Programs. Available <http://www.epa.gov/sites/production/files/2015-07/documents/lit-studies.pdf>. [accessed 17 January 2016].
- US EPA (U.S. Environmental Protection Agency). 2013. Materials Submitted to the National Research Council Part I: Status of Implementation of Recommendations by Integrated Risk Information System Program, January 30, 2013. Available: [http://www.epa.gov/sites/production/files/2014-06/documents/iris\\_program\\_materials\\_to\\_nrc\\_part\\_1.pdf](http://www.epa.gov/sites/production/files/2014-06/documents/iris_program_materials_to_nrc_part_1.pdf). [accessed 17 January 2016].
- US EPA (U.S. Environmental Protection Agency). 2014a. Framework for Human Health Risk Assessment to Inform Decision Making. Office of the Science Advisor. Risk Assessment Forum. EPA/100/R-14/001. Available: <http://www.epa.gov/risk/framework-human-health-risk-assessment-inform-decision-making>. [accessed 17 January 2016].
- US EPA (U.S. Environmental Protection Agency). 2014b. Next Generation Risk Assessment: Incorporation of Recent Advances in Molecular, Computational, and Systems Biology. EPA/600/R-14/004. Washington, DC. Available: <http://cfpub.epa.gov/ncea/risk/recordisplay.cfm?deid=286690#Download>. [accessed 17 January 2016].
- US FDA (U.S. Food and Drug Administration). 2009. Guidance for Industry: Evidence-Based Review System for the Scientific Evaluation of Health Claims – Final. Available: <http://www.fda.gov/Food/GuidanceRegulation/GuidanceDocumentsRegulatoryInformation/LabelingNutrition/ucm073332.htm>. [accessed 17 January 2016].

- US FDA (U.S. Food and Drug Administration). 2015. Plan to Increase Access to Results of FDA-Funded Scientific Research. Rockville, MD. Available: <http://www.fda.gov/downloads/ScienceResearch/AboutScienceResearchatFDA/UCM435418.pdf> [accessed 17 January 2016].
- U.S. House of Representatives. 1972. Federal Advisory Committee Act. Pub.L. 92-463, 6 October 1972. Available: [http://www.gsa.gov/graphics/ogp/without\\_annotations\\_R2G-b4T\\_0Z5RDZ-i34K-pR.pdf](http://www.gsa.gov/graphics/ogp/without_annotations_R2G-b4T_0Z5RDZ-i34K-pR.pdf) [accessed 17 January 2016].
- WHO IPCS (World Health Organization International Programme on Chemical Safety). 1999. Environmental Health Criteria 210. Principles for the Assessment of Risks to Human Health from Exposure to Chemicals. Geneva, Switzerland. Available: <http://www.inchem.org/documents/ehc/ehc/ehc210.htm>. [accessed 17 January 2016].
- WHO IPCS (World Health Organization International Programme on Chemical Safety). 2009. Environmental Health Criteria 240. Principles and methods for the risk assessment of chemicals in food. Geneva, Switzerland. Available: [http://www.inchem.org/documents/ehc/ehc/ehc240\\_index.htm](http://www.inchem.org/documents/ehc/ehc/ehc240_index.htm). [accessed 17 January 2016].
- WHO IPCS (World Health Organization International Programme on Chemical Safety). 2010. WHO Risk Assessment Tool Kit: Chemical Hazards. Geneva, Switzerland. Available: <http://www.who.int/ipcs/publications/methods/harmonization/toolkit.pdf?ua=1>. [accessed 17 January 2016].
- WHO (World Health Organization). 2012. WHO handbook for guideline development. Geneva, Switzerland. Available: [http://apps.who.int/iris/bitstream/10665/75146/1/9789241548441\\_eng.pdf?ua=1](http://apps.who.int/iris/bitstream/10665/75146/1/9789241548441_eng.pdf?ua=1). [accessed 17 January 2016].
- Woodruff TJ, Sutton P. 2014. The Navigation Guide systematic review methodology: a rigorous and transparent method for translating environmental health science into better health outcomes. *Environ. Health Perspect.* 122:1007–1014.

## Appendix 1. Guide for Judging the Quality of an Assessment

<b>Problem Formulation/Scoping and Planning</b>
<input type="checkbox"/> Possesses effective focus on identified risk management questions/options, and results in the depth/extent of the assessment are commensurate with the nature and significance of the decisions to be made.
<input type="checkbox"/> Issues pertaining to relevant exposure scenarios, data collection and evaluation, methods of assessment including mode of action analysis, ongoing research, etc. are identified.
<input type="checkbox"/> Evidence that a dialog occurred with scientific and stakeholder communities that afforded a reasonable opportunity for their input on key issues before (and while) preparing the assessment.
<b>Systematic Review of Literature</b>
<input type="checkbox"/> The search strategy (including predefined study inclusion/exclusion criteria, literature sources, search terms) used to identify relevant literature (both negative and positive studies) is well documented and is available to the public. Any restrictions placed on the literature search or data access are noted and explained.
<input type="checkbox"/> Evidence that outside parties were given a reasonable opportunity to provide relevant studies that were not identified in the authors' literature review and to comment on quality of studies selected for inclusion.
<input type="checkbox"/> Sufficient data for the critical studies and the models used in the assessment are available to interested external parties so as to enable them to replicate/verify the assessment outcomes and to judge the scientific credibility of the data/models.
<b>Hazard Assessment/Characterization</b>
<input type="checkbox"/> The quality ( <i>i.e.</i> , reliability, validity of method/study design), relevance, and utility of the

	results for each study are evaluated using an objective approach that employed pre-defined criteria.
<input type="checkbox"/>	An <i>a priori</i> established weight-of-evidence approach addressing causal relationships is applied in a systematic manner to integrate, weigh the lines of relevant evidence, and effectively use all relevant information. Both positive and negative studies are weighed objectively. Judgments and choices made are transparent.
<input type="checkbox"/>	There is a robust discussion of key lines of evidence and inherent uncertainties, alternative interpretations, other issues that may have prompted debate, and how these issues are addressed.
<input type="checkbox"/>	Questions of whether or not a response was adverse are identified, explained, and addressed.
<input type="checkbox"/>	Confounding factors or the extent to which other stressors might cause or impact the adverse effects ( <i>e.g.</i> , potential antagonistic/synergistic effects) with the subject chemical are considered and addressed.
<input type="checkbox"/>	Depending on available information and to the extent possible, mode of action (MoA) data are taken into account, evaluated in a systematic manner using pre-defined criteria, and are fully incorporated into the assessment of the key endpoints, dose-response relationships, human relevance, and life stage impacts.
<input type="checkbox"/>	The MoA analysis includes a consideration of category analogs as a complement to stressor-specific data, and existing knowledge is effectively leveraged on already established MoAs similar to the substance of interest.
<input type="checkbox"/>	There is a discussion of whether the key events within the MoA would progress to an adverse effect relative to concentration/dose and anticipated human exposure

(duration/magnitude/route), and life stage.
<input type="checkbox"/> Depending upon the purpose of the assessment, if the stressor produces an effect through a MoA by which other stressors have been shown to operate, the need for a cumulative assessment is identified.
<b>Dose-Response Assessment/Characterization</b>
<input type="checkbox"/> The endpoints used in the dose-response assessment are those most strongly causally associated with adverse responses, biologically plausible in humans, and derived from studies that meet standards of acceptable quality.
<input type="checkbox"/> The nature of responses ( <i>e.g.</i> , biochemical, morphological, physiological or functional change, severity of the effect, reversibility) and their dose-responses ( <i>e.g.</i> , steepness or shallowness of dose-response curve, dose spacing between NOAEL and LOAEL) are described.
<input type="checkbox"/> The dose responses are plotted for more than one endpoint of concern and a distribution of hazard values or points of departure (POD) are provided for all relevant endpoints. The selection of the hazard values is well justified and supported by the overall database.
<input type="checkbox"/> Consistent with the level of complexity needed and if data support modeling, multiple approaches are carried forward in the analysis and a justification is provided for model selection.
<input type="checkbox"/> Default assumptions are identified and the rationale for each is explained and their impact on the assessment's conclusions is described.
<input type="checkbox"/> What is known about endogenous production and naturally-occurring background levels of the subject chemical is considered, and if appropriate, incorporated into the analysis.
<input type="checkbox"/> Dose-response relationships are assessed for critical windows of exposure/susceptibility.

<input type="checkbox"/> Consistent with the level of complexity needed, and if available, suitable toxicokinetic and toxicodynamic data are used to derive more refined dose response estimates.
<input type="checkbox"/> Consistent with the level of complexity needed, and if available, quantitative dose-responses (and timing) of key events within a MOA are incorporated into the modeling.
<b>Exposure Assessment/Characterization</b>
<input type="checkbox"/> An understanding of the chemical's physiochemical properties, distribution and fate in the environment are reflected in the assessment.
<input type="checkbox"/> Chemical break down products are considered.
<input type="checkbox"/> Relevant populations are identified and assessed including demographic factors (e.g., life stage/age, sex, ethnicity), geography, and human activity patterns that would make a group more vulnerable to exposure than other groups.
<input type="checkbox"/> Depending on the purpose of the assessment, occupational-related activities are identified and assessed that bring workers into contact with the chemical.
<input type="checkbox"/> Intended uses, significant sources and scenarios of exposure (e.g., routes, frequency and durations) are identified and evaluated.
<input type="checkbox"/> The source of information used to derive the exposure estimates are described (e.g., from generic data, generic data with chemical-specific attributes, chemical-specific exposure monitoring data or internal dose data) and inherent uncertainties are identified.
<input type="checkbox"/> If models are used to estimate exposure, their strengths and limitations are clearly described and sufficient information is available to enable others to replicate and verify the modeling.
<input type="checkbox"/> If only conservative, worst-case estimates of exposure are generated, the rationale for the approach is provided.
<input type="checkbox"/> If the analysis is deterministic and did not employ a probabilistic approach, a justification is

provided.
<input type="checkbox"/> Sufficient, reliable data are used in lieu of default assumptions.
<input type="checkbox"/> Resulting certainties/uncertainties and default assumptions used are identified and a sensitivity analysis is conducted to evaluate the impact on conclusions.
<input type="checkbox"/> If only minimal information was available to assess exposure ( <i>e.g.</i> , physicochemical properties; molecular weight; vapor pressure solubility in fat and water), additional information needs are identified.
<b>Risk Characterization</b> (This phase summarizes the findings on hazard, dose response, and exposure characterizations, and develops an integrative analysis. The elements here should have also been addressed under the other three phases of risk assessment.)
<input type="checkbox"/> It is written for both technical and non-technical audiences and is clear and understandable in describing the purpose/objectives, scope, and main findings.
<input type="checkbox"/> Consistent with the scope and context, all potential hazards/risks are presented for the populations and exposure scenarios of interest.
<input type="checkbox"/> It reflects an appropriate matching of hazard and exposure data characterized by life stage and exposure scenario. If not, the assumptions used and their impact are described.
<input type="checkbox"/> It is consistent with data that meet the relevance and quality criteria, and reflects a minimization of bias on study design, data selection and interpretation, choice of models, and conclusions.
<input type="checkbox"/> Scientific facts and science policy choices are distinguished. Confidence in conclusions/risk values is placed clearly in the context of certainties and uncertainties, and the reasoning for use of and impact of defaults on conclusions are explained.

<input type="checkbox"/> If a quantitative uncertainty analysis is provided, it is probabilistic and the data, methods, and models used are described to allow for independent re-analysis. If not, a justification for not doing a quantitative analysis is provided.
<input type="checkbox"/> Variability in effect or response across relevant populations(s) is discussed with significant uncertainties noted.
<input type="checkbox"/> Alternative judgments, hypotheses and models are presented along with support underlying these alternatives. If the assessment includes only a worst-case scenario, an explanation is provided.
<input type="checkbox"/> Significant data needs are clearly identified. There is discussion of the potential impact such data might have on the assessment ( <i>i.e.</i> , value of information).
<input type="checkbox"/> If appropriate, risk-risk comparisons are included to provide context for the decision-maker.
<b>Peer Review</b>
<input type="checkbox"/> A documented process for peer review consistent with the agency's/sponsor's guidance/policies and its extent and nature matches the purpose/scope and potential impact of the assessment. It is conducted prior to the assessment being finalized.
<input type="checkbox"/> Conflicts of interest and bias are identified and addressed.
<input type="checkbox"/> All draft materials are available to public commenters and peer reviewers at a similar time and adequate time is allowed for public comment.
<input type="checkbox"/> Peer reviewers receive the public comments in advance for adequate consideration before the peer review meeting is conducted.
<input type="checkbox"/> There is a reasonable opportunity for public comments to be presented at the public peer review meeting, and there is an opportunity for peer reviewers to engage with public commenters on the key technical issues they put forward.

<input type="checkbox"/> If peer reviewers did not reach consensus, a minority opinion/report is provided.
<input type="checkbox"/> Public and peer review comments are objectively and appropriately addressed.